



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification:</b> <b>C12N 15/52, C12N 5/10,</b> <b>C12N 9/00, C12N 15/63,</b> <b>C12P 17/06</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/22139</b> <b>(43) International Publication Date:</b> 20 April 2000 (20.04.2000)
<b>(21) International Application Number:</b> PCT/US99/23535 <b>(22) International Filing Date:</b> 11 October 1999 (11.10.1999) <b>(30) Priority Data:</b> 198 46 493.2 09 October 1998 (09.10.1998) DE <b>(60) Parent Application or Grant</b> GESELLSCHAFT FÜR BIOTECHNOLOGISCHE FORSCHUNG MBH (GBF) [/]; (). BRISTOL-MYERS SQUIBB COMPANY [/]; (). BEYER, Stefan [/]; (). BLOECKER, Helmut [/]; (). BRANDT, Petra [/]; (). CINO, Paul, M. [/]; (). DOUGHERTY, Brian, A. [/]; (). GOLDBERG, Steven, L. [/]; (). HOFLE, Gerhard [/]; (). MUELLER, Rolf- Joachim [/]; (). REICHENBACH, Hans [/]; (). BEYER, Stefan [/]; (). BLOECKER, Helmut [/]; (). BRANDT, Petra [/]; (). CINO, Paul, M. [/]; (). DOUGHERTY, Brian, A. [/]; (). GOLDBERG, Steven, L. [/]; (). HOFLE, Gerhard [/]; (). MUELLER, Rolf-Joachim [/]; (). REICHENBACH, Hans [/]; (). SANTUCCI, Ronald, R. ; ().		<b>Published</b>
<b>(54) Title: DNA SEQUENCES FOR ENZYMATIC SYNTHESIS OF POLYKETIDE OR HETEROPOLYKETIDE COMPOUNDS</b> <b>(54) Titre: SEQUENCES D'ADN DESTINEES A LA SYNTHÈSE ENZYMATIQUE DE COMPOSÉS À BASE DE POLYKETIDES</b> <b>OU D'HETEROPOLYKETIDES</b>  <b>(57) Abstract</b> <p>The invention consists of: (1) cloned <i>Sorangium cellulosum</i> polyketide synthase (PKS) biosynthetic cluster DNA; and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to: (a) increasing yields of PKS product in <i>Sorangium cellulosum</i> (e.g., by amplification or genetic modification of the epothilone gene cluster or its component parts); (b) increasing yields of polyketide product in a heterologous system by transfer of the epothilone gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS gene cluster or its component parts; (c) modification of the polyketide product chemical structure in either <i>Sorangium cellulosum</i> or a heterologous host (e.g., by genetic modification of the epothilone gene cluster or its component parts; and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmids and plasmids: A2 cosmid; the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7); pEPOcos8 cosmid; A5 cosmid; Sau4 (10 kb plasmid).</p> <b>(57) Abrégé</b> <p>L'invention concerne: (1) un ADN biosynthétique cloné en grappe de polykétide synthase (PKS) de <i>Sorangium cellulosum</i>; et (2) la séquence nucléotidique et les séquences protéiques codantes prévues de l'ADN cloné. L'invention peut avoir les applications suivantes (sans caractère limitatif): (a) augmentation de la production de PKS chez <i>Sorangium cellulosum</i> (p.ex., par l'amplification ou la modification génétique de la grappe de gènes épothilone ou de ces parties constitutives); (b) augmentation de la production du produit polykétide dans un système hétérologue par le transfert de la grappe de gènes épothilone ou de ces parties constitutives, qui peut être suivie par l'amplification ou la modification génétique de la grappe de gènes PKS ou des ses parties constitutives; (c) modification de la structure chimique du produit polykétide soit chez <i>Sorangium cellulosum</i> soit chez un hôte hétérologue (p.ex., par l'amplification ou la modification génétique de la grappe de gènes épothilone ou de ces parties constitutives); et (d) détection de gènes et de produits géniques participant à la fabrication de polykétides ou de molécules correspondantes dans d'autres organismes (p.ex., par des dosages à hybridation ou à complémentation). La séquence d'ADN et l'analyse sont présentées pour les cosmides et les plasmides suivants: cosmide A2; région pEPOcos6 (se chevauchant avec pEPOcos6 et pEPOcos7); cosmide pEPOcos8; cosmide A5; Sau4 (plasmide 10 kb).</p>		

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>C12N 15/52, 9/00, 15/63, 5/10, C12P 17/06</b>	<b>A2</b>	(11) International Publication Number: <b>WO 00/22139</b>
		(43) International Publication Date: <b>20 April 2000 (20.04.00)</b>

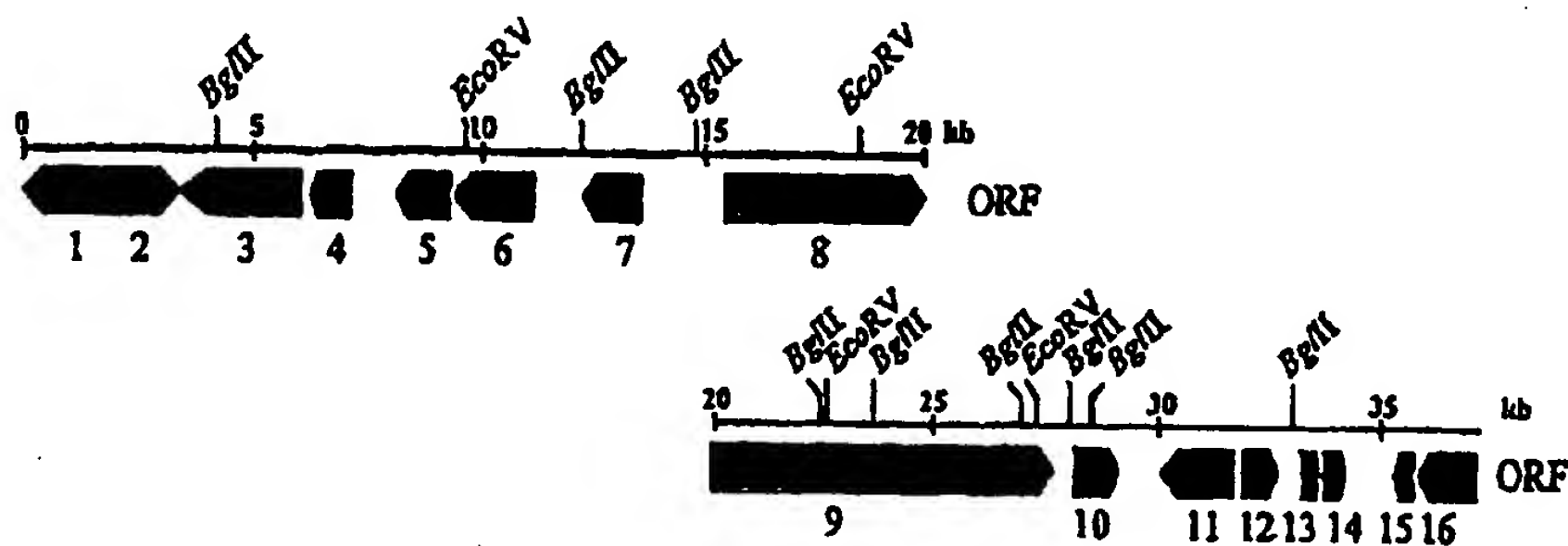
(21) International Application Number: **PCT/US99/23535**(22) International Filing Date: **11 October 1999 (11.10.99)**(30) Priority Data:  
**198 46 493.2      9 October 1998 (09.10.98)      DE**(71) Applicants (for all designated States except US):  
**GESELLSCHAFT FÜR BIOTECHNOLOGISCHE FORSCHUNG MBH (GBF) [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BRISTOL-MYERS SQUIBB COMPANY [US/US]; P.O. Box 4000, Princeton, NJ 08543-4000 (US).**

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BEYER, Stefan [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BLOECKER, Helmut [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BRANDT, Petra [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). CINO, Paul, M. [US/US]; 4 Crest Drive, Bound Brook, NJ 08805 (US). DOUGHERTY, Brian, A. [US/US]; 10 Rosemary Lane, Killingworth, CT 06419 (US). GOLDBERG, Steven, L. [US/US]; 25 Compton Court, Basking Ridge, NJ 07920 (US). HOFLE, Gerhard [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). MUELLER, Rolf-Joachim**[DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). **REICHENBACH, Hans [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE).**(74) Agents: **SANTUCCI, Ronald, R. et al.; Kane, Dalsimer, Sullivan and Levy, LLP 20th floor, 711 Third Avenue, New York, NY 10017 (US).**(81) Designated States: **AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).**

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: **DNA SEQUENCES FOR ENZYMATIC SYNTHESIS OF POLYKETIDE OR HETEROPOLYKETIDE COMPOUNDS**

(57) Abstract

The invention consists of: (1) cloned *Sorangium cellulosum* polyketide synthase (PKS) biosynthetic cluster DNA; and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to: (a) increasing yields of PKS product in *Sorangium cellulosum* (e.g., by amplification or genetic modification of the epothilone gene cluster or its component parts); (b) increasing yields of polyketide product in a heterologous system by transfer of the epothilone gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS gene cluster or its component parts; (c) modification of the polyketide product chemical structure in either *Sorangium cellulosum* or a heterologous host (e.g., by genetic modification of the epothilone gene cluster or its component parts; and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmid and plasmids: A2 cosmid; the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7); pEPOcos8 cosmid; A5 cosmid; Sau4 (10 kb plasmid).

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## D scription

5

10

15

20

25

30

35

40

45

50

55



5

10

15

DNA sequences for enzymatic synthesis of polyketide or  
heteropolyketide compounds

10

20

The present invention relates to DNA sequences for enzymatic synthesis of polyketide or heteropolyketide compounds produced by the bacterium *Sorangium cellulosum*.

25

15 Background and introduction

30

20

35

40

45

30

This patent application describes DNA sequences for the enzymatic synthesis of polyketide and/or heteropolyketide structures synthesized by the myxobacterium *Sorangium cellulosum*. Several of these compounds have known cytotoxic, immunosuppressive, antibiotic and fungicidal biological activity, with the epothilones having been most studied and characterized. The fermentation of large quantities of secondary metabolites from microorganisms, especially from myxobacteria, is a time consuming and difficult process that often involves complications (e.g. contamination, low product yield, difficult isolation and purification). Therefore it would be advantageous to use a well-characterized organism for such fermentations. After cloning of the desired biosynthetic genes one could create such an organism via genetic engineering and manipulate the biosynthesis of the compound. Identified sequences

50

55

5

2

can be cloned into optimized expression vectors and generate recombinant cell lines that overproduce polyketide structures.

10

Polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) represent macromolecular and multifunctional enzymes which are characterized by a modular architecture. PKS condenses activated carbonic acids (usually acetate and propionate) and reduce the resulting 2-keto acid intermediates stepwise in a fatty acid biosynthesis-like fashion. Responsible for each reaction step is a specific domain that recognizes, activates, condenses and reduces the carbonic acid. Depending on the presence of these domains in the corresponding modules, every reduction stage can occur in the final product (Rawlings, Nat. Prod. Reports 14, 523-556 [1997]; for a review, see Chem. Rev. 97, 2463-2760 [1997]). A typical example for the biosynthesis of a polyketide is the macrolide antibiotic erythromycin (Staunton and Wilkinson, Chem. Rev. 97, 2611-2630 [1997]). NRPSs are also modular enzymes and condense via peptide bonds amino acids to low molecular weight bioactive substances like bacitracin or tyrocidin. Typical domains of these systems activate the amino acid and condense it with the growing peptide chain. Methylations, epimerisations and modifications via additional protein domains are possible (Stachelhaus and Marahiel, FEMS Microbiol Lett. 125, 3-14 [1995]). Both types of enzymes (NRPS and PKS) share the modular organization of the proteins in which specific catalytic domains are responsible for recognition, activation, condensation and modification of the single elongation units. The growing chain of amino acids and/or carbonic acids is extended through the action of one module adding one unit. The domains of each module carry the active centers responsible for the enzymatic steps of the biosynthesis.

5

15

20

25

30

35

40

45

50

55

5

3

10

15

20

25

30

35

40

45

30

50

55

Little is known about the biosynthesis of biologically active polyketides and polypeptides from myxobacteria. Fragments of the biosynthetic gene clusters of soraphen and saframycin have been described from *Sorangium cellulosum* So ce26 and *Myxococcus xanthus*, respectively (Schupp et al., J. Bacteriol. 177, 3673-3679 [1995] and Pospiech et al., Microbiology 141, 1793-1803 [1995]). We have constructed genomic libraries of the epothilone producer *Sorangium cellulosum* So ce90. Gene probes based on PKS and PS genes were used to isolate recombinant cos-mids, which were then sequenced and characterized. Several unique pathways containing PKS, PS, or a combination of both types of genes were identified, demonstrating that this organism is potentially a rich source of novel bioactive compounds.

A subject of the present invention is therefore to provide DNA sequences according to claim 1 the expression products of which perform or are involved in the enzymatic biosynthesis, mutasynthesis or partial synthesis of polyketide or heteropolyketide compounds. The DNA sequences may be inserted into well known and optimized expression vectors by common techniques of molecular biology, thus allowing transformation, selection and cloning of cells, which cells are then capable of synthesizing polyketide or heteropolyketide compounds by fermentation. Using an overproducing clone allows the desired polyketide or heteropolyketide compounds be easily produced and recovered in high amounts. Further, knowledge of the localization of regulatory DNA segments and individual structural genes allows "site-directed mutagenesis" using common techniques for genetic engineering, and thus construction of optimized enzymes ("protein engineering") for fermentative synthesis of polyketide or heteropolyketide compounds.

5

4

10

The invention thus further relates to a recombinant expression vector according to claim 16, cells transformed therewith according to claim 17 and to a process for enzymatic biosynthesis, mutasynthesis or partial synthesis of polyketide or heteropolyketide compounds according to claim 23.

15

Preferred and/or advantageous embodiments of the present invention are subject-matter of the subclaims.

20

In brief, the invention consists of (1) cloned *Sorangium cellulosum* polyketide synthase (PKS) and/or peptide synthetase (PS) biosynthetic cluster DNA and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to, (a) increasing yields of PKS product in *Sorangium cellulosum* (e.g., by amplification or genetic modification of the epothilone gene cluster or its component parts), (b) increasing yields of polyketide and/or peptide synthetase product in a heterologous system by transfer of the corresponding gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS and/or PS gene cluster or its component parts, (c) modification of the polyketide and/or peptide synthetase product chemical structure in either *Sorangium cellulosum* or a heterologous host (e.g., by genetic modification of the corresponding gene cluster or its component parts) and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmids and plasmids:

45

- A2 cosmid as defined in claim 6
- the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7) as defined in claim 7

30

50

55

5

5

10

- pEPOcos8 cosmid as defined in claim 10
- A5 cosmid as defined in claim 12
- Sau4 (10 kb plasmid) as defined in claim 14

15

5 The invention is now described in more detail by examples and for illustration only. The examples are not to be construed as any limitation of the scope.

20

10 Figure 1 is a restriction map of one of the DNA sequences of the present invention (cosmid A2 insert) indicating also the localization of regulatory DNA segments and the individual structural genes ("open reading frames" or ORFs) 1 to 16.

25

15 Figure 2 shows the open reading frames found on pEPOcos6 region

30

DNA sequence data from A2 cosmid are as defined in claim 6.

35

Table 1 correlates ORFs 1 to 16 found on A2 cosmid with the respective biological function (Regulators, Enzymes).

40

45

50

55

5

6

Table 1

10

	gene/function	position
ORF 1	regulatory element	1666 - 1
ORF 2	regulatory element	1605 - 3338
ORF 3	acyl-t-RNA synthetase	6100 - 3398
ORF 4	monooxygenase	7110 - 6374
ORF 5	amino transferase	9590 - 8433
ORF 6	L-dopa decarboxylase	11393 - 9855
ORF 7	oxidoreductase	13656 - 12712
ORF 8	polyketide synthase	15374 - 18984
ORF 9	polypeptide synthetase	20003 - 27889
ORF 10	peptidase	28251 - 29402
ORF 11	regulatory element	31720 - 30401
ORF 12	sigma factor	31982 - 32932
ORF 13	regulatory element	33128 - 33613
ORF 14	regulatory element	33661 - 34007
ORF 15	transcription regulator	35611 - 35255
ORF 16	signal transduction	37856 - 35730

15

20

25

30

35

40

45

50

55

5

7

## Working Examples

10

A. Construction of a *Sorangium cellulosum* cosmid library

5

15

1. Isolation of genomic DNA from *S. cellulosum* So ce90

20

25

30

35

40

25

45

30

50

55

5

8

10

with 70% and 95% ethanol, air-dried, and resuspended in 0.5 ml TE buffer (0.01 M Tris-HCl, 0.001 M ethylenediaminetetraacetic acid [EDTA], pH 8.0).

15

20

5 b. Alternatively, genomic DNA was isolated from *S. cellulosum* cells cultured as described in section A.1 using the Midi Qiagen Blood & Cell Culture DNA purification Kit (Qiagen, Hilden, Germany) following the Qiagen Genomic DNA Handbook protocol for bacterial DNA isolation (1997, Qiagen, Hilden, Germany, p. 29 ff.). In order to obtain high molecular weight chromosomal DNA the precipitated DNA was recovered with a bent pasteur pipette as described in section A.1.

25

## 15 2. Isolation of plasmid DNA

30

35

40

45

a. pFD666: pFD666 is a bifunctional *E. coli*-*Streptomyces* cosmid cloning vector (see Denis and Brzezinski, Gene 111, 115-118 [1992]). To maintain stability of large inserts, it is present in low-medium copy number when replicated in *E. coli*. For this reason, isolation of sufficient pure DNA to carry out cloning experiments was difficult using commercial kits with standard protocols. A modified procedure was therefore used to obtain pFD666 DNA. A 10 ml culture of DH10B(pFD666) was grown for 16-20 hr at 37°C in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.0) medium containing 50 µg/ml kanamycin sulfate. Fifty ml of LB + kanamycin was inoculated to a starting OD<sub>600</sub> of ca. 0.25 and shaken at 300 rpm, 37°C, until the OD<sub>600</sub> reached ca. 0.6. Five hundred ml of LB + kanamycin medium in a 2 l flask was inoculated with 25 ml of this culture and incubated

50

55



5

9

10

15

20

25

30

35

40

45

50

55

under the same conditions for 2.5 hr. Chloramphenicol ( 2.5 ml of a 34 mg/ml solution in 100% EtOH) was added and the incubation continued for an additional 16-20 hr. (The previous steps were performed according to Maniatis et al. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.) Cells were pelleted for 10 min, 16,000 x g . They were resuspended in 9 ml of 50 mM glucose/25 mM Tris-HCl (pH 8.0)/10 mM EDTA and transferred to a 50 ml disposable centrifuge tube. One ml of a freshly-prepared 10 mg/ml lysozyme solution in 10 mM Tris-HCl, pH 8.0 was added and the cell suspension incubated in a 37°C water bath for 10 min. Twenty ml of a freshly-prepared 0.2 NaOH/1% sodium dodecyl sulfate (SDS) solution was added and the tube inverted gently 5-7 times to mix the contents. After 5 min at room temperature, 15 ml of 5 M potassium acetate (pH 4.8) was added and the tube inverted sharply 3-4 times. The tube was centrifuged at 6,000 x g for 10 min at 4°C and the supernatant poured through 2 layers of sterile cheese cloth into a fresh 50 ml disposable tube. Isopropanol to a final concentration of 0.6% was added and the contents of the tube mixed several times. The precipitated nucleic acid was centrifuged at 6,000 x g for 10 min at 4°C. The pellet was washed with 70% EtOH and any excess EtOH was aspirated from the pellet, which was allowed to air dry for 5 min. It was resuspended in 5 ml of 50 mM 3-(N-Morpholino)propanesulfonic acid (MOPS)/750 mM NaCl, pH 7.0 and added to an equilibrated QIAfilter Midi column (Qiagen, Chatsworth, CA). The manufacturer's protocol for washing and eluting the plasmid DNA was followed.

5

10

10

b. SuperCos: SuperCos plasmid DNA was purchased from Stratagene (La Jolla, CA).

5

### 3. Preparation of ca. 38-47 kb *Sau3A1* fragments of *S. cellulosum* chromosomal DNA

15

20

10

25

15

30

35

40

25

45

30

50

55

a. *S. cellulosum* chromosomal DNA prepared as described in section A.1.a was partially cleaved with restriction endonuclease *Sau3A1* in a 1000  $\mu$ l reaction volume consisting of 50  $\mu$ g chromosomal DNA, 5 units enzyme (Promega, Madison, WI), 0.006 M Tris-HCl, 0.006 M MgCl<sub>2</sub>, 0.10 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 5 min at 37°C. The reaction mixture was extracted once with an equal volume of 1:1 phenol:chloroform. After centrifugation, the upper aqueous phase was saved, to which 0.1 vol. of 3 M sodium acetate and 0.6 vol. isopropanol was added. DNA was pelleted by centrifugation for 5 min at 16,000 x g in a microfuge and washed once with 0.5 ml 70% EtOH. After drying in a SpeedVac (Savant Instruments, Farmingdale, NY) for 5 min, the pellet was resuspended in 0.1 ml TE buffer. The DNA was layered on top of a 12 ml 10-40% sucrose gradient prepared in TE buffer and centrifuged at 113,600 x g for 16 hr, 10°C using a Beckman SW40Ti rotor (Beckman Instruments, Palo Alto, CA). Five hundred  $\mu$ l aliquots of the gradient were removed using a pipetor beginning at the top of the tube. Samples (5  $\mu$ l) of the fractions were analyzed by electrophoresis through a 0.5% agarose gel in TAE buffer (0.04 M Trizma base, 0.02 M acetic acid, and 0.001 M EDTA, pH 8.3) containing 0.5  $\mu$ g/ml ethidium bromide for 6 hr at 100 V. Fractions containing DNA fragments of ca. 40-45 kb were identified by comparison to a high molecular weight DNA standard (Life Technologies,

5

10

15

20

25

30

35

40

45

50

55

Gaithersburg, MD). Sucrose was diluted from the corresponding 0.5 ml fraction by addition of 0.5 vol. TE. Subsequently, DNA was precipitated by addition of 0.1 vol. 3 M sodium acetate and 0.6 vol. isopropanol. DNA was pelleted by centrifugation at 16,000 x g for 10 min in a microfuge. DNA was washed with 0.5 ml 70% EtOH and dried in a SpeedVac with moderate heat for 10 min. Finally, the DNA was resuspended in distilled H<sub>2</sub>O at a concentration of 0.5 mg/ml.

b. Alternatively, 10 µg of *S. cellulosum* chromosomal DNA prepared as described in A.1.b was treated with 0.3 U Sau3A1 (New England Biolabs, Beverly, MA) for 1 h at 37°C in 400 µl of the supplier's recommended reaction buffer. Formation of DNA fragments of about 40 kb in size was checked by comparison of the motility behavior with high molecular weight DNA standards after a 0.3% agarose gel electrophoresis. An equal volume of phenol:chloroform (1:1) was added, mixed and centrifuged. The upper aqueous phase was recovered and 0.1 vol. of 3 M sodium acetate and 0.6 vol. of isopropanol were added. After centrifugation, the precipitated DNA was washed twice with 0.5 ml 70% ice cold ethanol and finally air-dried. The DNA fragments were resuspended in 100 µl shrimp alkaline phosphatase reaction buffer and dephosphorylated for 150 min. at 37°C using 2 U shrimp alkaline phosphatase (Amersham Life Science, Cleveland, OH). A phenol:chloroform extraction followed as described above. Finally, the DNA was precipitated by addition of 0.1 vol. 3 M sodium acetate and 0.6 vol. isopropanol, dried, and dissolved in TE buffer.

30

5

12

#### 4. Preparation of cosmid libraries

10

a. Using pFD666: Vector pFD666 was cleaved with restriction endonuclease BamHI in a 0.02 ml reaction volume consisting of 2  $\mu$ g plasmid DNA, 10 units of BamHI (Promega), 0.006 M Tris-HCl, 0.006 M MgCl<sub>2</sub>, 0.05 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 90 min at 37°C. Five  $\mu$ l of 10x alkaline phosphatase buffer (0.5 M Tris-HCl [pH 9.3], 0.01 M MgCl<sub>2</sub>, 0.001 M ZnCl<sub>2</sub>, 0.01 M spermidine) was added to the reaction followed by alkaline phosphatase (0.01 units/pmol ends; Promega) and distilled H<sub>2</sub>O to a final volume of 0.05 ml. The sample was incubated for 30 min at 37°C and a second aliquot of phosphatase was added. After a further 30 min at 37°C, 0.3 ml of stop buffer (0.01 M Tris-HCl [pH 7.5], 0.001 M EDTA, 0.2 M NaCl, 0.5% SDS) and 0.35 ml of 1:1 phenol; CHCl<sub>3</sub> was added to the reaction. The sample was mixed gently several times by inversion and centrifuged at 16,000 x g for 3 min to separate the phases. The aqueous layer was removed to a new microfuge tube. 0.1 vol. 3 M sodium acetate and 2 vol. 100% EtOH were added and the precipitated DNA pelleted by centrifugation at 16,000 x g for 10 min. Liquid was removed by aspiration and the pellet washed once with 0.5 ml 70% EtOH. The DNA was dried in a SpeedVac and resuspended in TE buffer to 0.5 mg/ml.

40

Digested, phosphatase-treated pFD666 was ligated to the partially-cleaved chromosomal DNA (see sections A.3.a and B.1.a) in a 0.005 ml reaction consisting of 1  $\mu$ g pFD666, 1  $\mu$ g *S. cellulosum* DNA, 0.03 M Tris-HCl (pH 7.8), 0.01 M MgCl<sub>2</sub>, 0.01 M dithiothreitol, and 0.0005 M adenosine-5'-triphosphate and 1.5 Weiss units of T4 DNA ligase (Promega). The reaction was carried out at room temperature for 2 hr. The entire reaction

50

55

5

13

10

15

20

10

25

15

30

35

20

40

25

45

30

50

55

mix was packaged into bacteriophage  $\lambda$  *in vitro* using Packagene extracts (Promega) according to the manufacturer's directions. The entire packaging reaction (0.5 ml) was diluted with 4.5 ml SM buffer (per liter: 5.8 g NaCl, 2 g  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 1 M Tris-HCl [pH 7.5], 5 ml 2% gelatin solution). Transfection was performed by adding 10 ml of an overnight culture of *E. coli* DH5 $\alpha$  that had been grown in LB medium with 0.01 M  $\text{MgSO}_4$  and 0.2% maltose to the diluted phage and incubating at 37°C for 20 min. 0.8 ml of LB was added and the cells shaken at 225 rpm for 1 hr at 37°C. Cells were pelleted, resuspended in LB, and spread onto a 150 mm LB + kanamycin agar plate. After 3 d. at 30°C, the colonies were harvested by picking ca. 800 colonies into 2.0 ml LB + kanamycin medium containing 20% glycerol, freezing on dry ice, and storing at -70°C. In addition, six kanamycin-resistant colonies were inoculated into 2 ml LB + kanamycin liquid medium and incubated at 37°C, 250 rpm, for 18-24 hr. Cosmid DNA was prepared using a standard alkaline lysis procedure starting with 1.5 ml of the culture. DNA was digested with restriction endonuclease PstI and samples electrophoresed on a 0.8% TAE agarose gel for 1.5 hr at 100 V. A unique restriction pattern was noted in each sample and the total size of the insert was calculated to be between 40 and 45 kilobases.

**b. Using SuperCos:** 30  $\mu\text{g}$  of vector SuperCos was digested with XbaI (New England Biolabs, Beverly, MA) for 210 min at 37°C in 100  $\mu\text{l}$  of the recommended reaction buffer. Ten  $\mu\text{l}$  sodium acetate and 60  $\mu\text{l}$  isopropanol was added before the solution was centrifuged for 30 min at 16,000 x g. The precipitated DNA was washed twice with 500  $\mu\text{l}$  ice cold 70% ethanol. The vector DNA was precipitated and air-dried, dissolved in 135  $\mu\text{l}$  shrimp al-

5

14

10

15

20

kaline phosphatase reaction buffer and treated with 2.5 U shrimp alkaline phosphatase for 150 min. After heat inactivation of the enzyme at 75°C for 20 min, a phenol:chloroform extraction was performed as described in section 1. c. The DNA, resuspended in 100 µl BamHI restriction buffer was hydrolyzed with 15 U BamHI (New England Biolabs, Beverly, MA) for 180 min. A phenol:chloroform extraction followed (see section A.3). The SuperCos DNA was precipitated by addition of 0.1 vol 3 M sodium acetate and 0.6 vol isopropanol, centrifuged at 16,000 x g, and resuspended in 50 µl TE buffer.

25

15

30

35

40

45

Four µg of digested vector DNA was ligated with 10 µg partially hydrolyzed genomic DNA from *S. cellulosum* (as described in section A.3.b) in a final volume of 20 µl using 2 U T4 DNA ligase and the appropriate reaction buffer (Gibco BRL, Eggenstein, Germany). The reaction was carried out at 16°C overnight. The reaction mixture was packaged into phage particles using the Gigapack III XL packaging extract kit (Stratagene) according to the manufacturer's protocol. Treatment of packaging reaction mixture and transfection of *E. coli* SURE (Stratagene) was performed as described in 4.a. Transfected cells were concentrated by centrifugation, resuspended in fresh LB medium and distributed on LB agar plates containing 50 µg/ml<sup>-1</sup> kanamycin. The plates were incubated overnight at 30°C. 1600 recombinant clones were transferred into 96 well microtiter plates filled with 80 µl LB medium containing 50 µg/ml<sup>-1</sup> kanamycin per well and propagated overnight at 30°C. The following day the microtiter plates were used to inoculate a second set of microtiter plates in order to obtain a duplicate of the recombinant clones. Each well of the original set of microtiter plates was supplemented with 80 µl 50 % glycerol and the entire plate stored at -70°C.

50

55

5

15

10

15

20 randomly chosen transformants were inoculated into 3 ml LB medium with 50  $\mu\text{g}/\text{ml}^{-1}$  kanamycin and incubated over night at 37°C in order to isolate plasmid DNA using the Qiagen plasmid extraction kit (Qiagen, Hilden, Germany). Restriction fragment analysis of the recombinant cosmids using the restriction endonucleases PstI and BglII indicated that the cosmids contained inserts of approximately 35 to 42 kb in size.

20

10

#### B. Construction of a *S. cellulosum* plasmid library

25

30

35

40

##### 1. Preparation of 8-12 kb fragments of *S. cellulosum* chromosomal DNA.

*S. cellulosum* chromosomal DNA prepared as described in section A.1.a was partially cleaved with restriction endonuclease Sau3A1 in a 100  $\mu\text{L}$  reaction volume consisting of 5  $\mu\text{g}$  chromosomal DNA, 5 units enzyme (Promega, Madison, WI), 0.006 M Tris-HCl, 0.006 M  $\text{MgCl}_2$ , 0.10 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 4 min at 37°C. The digested DNA was electrophoresed through a 11 x 14 cm 0.8% TAE-agarose gel for 18 hr at 17 V. Fragments of 8-12 kb were cut from the gel and purified using the QIAquick Gel Extraction Kit using the manufacturer's protocol (Qiagen).

##### 2. Preparation of the plasmid library

45

Plasmid pZero2.1 (Invitrogen, Carlsbad, CA) was cleaved with restriction endonuclease BamHI in a 0.02 ml reaction volume consisting of 1  $\mu\text{g}$  plasmid DNA, 10 units of BamHI (Promega), 0.006 M Tris-HCl, 0.006 M  $\text{MgCl}_2$ , 0.05 M NaCl, and 0.001 M di-

50

55

thiothreitol ( pH 7.5) for 20 min at 37°C. 0.08 ml of dH<sub>2</sub>O and 0.1 ml of 1:1 phenol:CHCl<sub>3</sub> was added. The sample was briefly vortexed and centrifuged at 16,000 x g for 2 min. The aqueous layer was removed to a new microfuge tube. 0.1 vol. 3 M sodium acetate and 2 vol. 100% EtOH were added and the precipitated DNA pelleted by centrifugation at 16,000 x g for 10 min. Liquid was removed by aspiration and the pellet washed once with 0.5 ml 70% EtOH. The DNA was dried in a SpeedVac and resuspended in TE buffer to 0.004 µg/ml. Digested pZero2.1 was ligated to the partially-cleaved chromosomal DNA in a 0.01 ml reaction consisting of 0.004 µg pZero2.1, 0.05 µg *S. cellulosum* DNA, 0.03 M Tris-HCl (pH 7.8), 0.01 M MgCl<sub>2</sub>, 0.01 M dithiothreitol, and 0.0005 M adenosine-5'-triphosphate and 1.5 Weiss units of T4 DNA ligase (Promega). The reaction was carried out at room temperature for 2 hr. 0.015 ml dH<sub>2</sub>O and 0.25 ml of 1-butanol were added, the sample vortexed briefly, and centrifuged at 16,000 x g for 10 min. Liquid was aspirated away from the pellet and the sample dried in a SpeedVac for 5 min. The ligated DNA was resuspended in 0.005 ml dH<sub>2</sub>O and mixed with 0.04 ml of electro-competent *Escherichia coli* DH10B cells (GIBCO/BRL, Gaithersburg, MD). The sample was placed into a pre-chilled 0.2 mm-gap electroporation cuvette and transformed into the bacteria by electroporation using a BioRad Gene Pulser II unit (BioRad, Hercules, CA) at 25 µF and 200 Ω. 0.96 ml SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 20 mM MgSO<sub>4</sub>, 20 mM glucose) was mixed with the cells and transferred to a 1.5 ml microfuge tube. The sample was incubated at 37°C, 225 rpm, for 1 hr. Aliquots of the cells were spread onto an LB agar + kanamycin and incubated at 37°C for 20 hr to estimate the number of transformants obtained. Six kana-



mycin resistant colonies were confirmed to contain an insert of the expected size as described in section A.4.a.

### C. Identification of cosmids possessing polyketide synthase genes

#### 1. Colony blot hybridizations using cosmid library in pFD666:

A 20 x 20 cm sheet of Duralon UV membrane (Stratagene) was placed on top of a 24.5 x 24.5 cm square bioassay dish containing 250 ml LB agar - kanamycin. An aliquot of the frozen cosmid library in 1 ml LB medium was spread on the filter. The plate was incubated at 37°C for 24 hr. Colonies were replicated onto two fresh filters which were placed onto LB + kanamycin agar medium and incubated at 28°C for 18 hr. Lysis of cells and neutralization of released DNA was performed according to directions that were provided with the filters. The DNA was crosslinked to the filters using a UV Stratalinker 2400 unit (Stratagene) in the auto crosslink mode. Cell debris was removed by placing the filters in a container with a solution of 3 X SSC (20 X SSC contains, per liter, 173.5 g NaCl, 88.2 g sodium citrate, pH adjusted to 7.0 with 10 N NaOH), 0.1% SDS and rubbing the lysed colonies with a Kimwipe. The filters were then incubated at least 3 hr with the same wash solution for at least 3 hr at 65°C. The plasmid library was treated similarly except cells were spread onto a 137 mm circular Duralon UV membrane placed on top of a 150 mm petri dish containing 80 ml LB agar + kanamycin.

For hybridizations, a probe consisting of a 650-base pair (bp) polymerase chain (PCR) fragment representing a portion of a *S. cellulosum* polyketide synthase gene was used. The fragment

was amplified using primers to consensus regions of Type I (macrolide) polyketide synthase (PKS) genes (Swan et al., Mol. Gen. Genetics 242, 358-362 [1994]). A series of sense and anti-sense oligonucleotides were prepared for PCR studies as indicated in the following table 2:

Table 2

Oligo-nucleotide	I. DNA sequence (5' → 3')	Corresponding amino acid sequence
120 (sense)	CGGT (C/G) AAGTC (C/G) AACATCGG	KSNIGHT
121 (anti-sense)	GC (A/G) ATCTC (A/G) CCCTGCCA (A/G) TG	HSQGEIA
122 (sense)	GT (C/G) GACAC (C/G) GC (C/G) TGCTC (C/G)	VDTACSS
123 (sense)	GG (C/G) AC (C/G) AACGC (C/G) CACGT (C/G) A T	GTNAHVI
124 (anti-sense)	CCCTG (C/G) CC (C/G) GGGAA (C/G) ACGAA	FVFPQGQ

10

The selection of C or G where necessary in the third position of a codon reflects the very high overall G + C content of *S. cellulosum* (ca. 70%). Conditions for PCR were as follows: 0.01 M Tris-HCl (pH 9.0), 0.05 M KCl, 0.003 M MgCl<sub>2</sub>, 0.1% Triton X-100, 200 μM of each primer, 2.5 U Taq DNA polymerase (Promega), 5.0% dimethyl sulfoxide (Sigma), and 0.01 μg of *S. cellulosum* chromosomal DNA in a 0.05 ml reaction volume. Reactions were

50

55

5

19

10

15

20

carried out in a Perkin-Elmer Model 480 Thermocycler (Perkin-Elmer Corporation, Foster City, CA) under the following conditions: 94°C, 1 min; 50°C, 1 min, 72°C, 1.5 min for a total of 30 cycles. Each possible combination of sense and anti-sense primers were tried. A 650-bp and 350-bp fragment was amplified using oligos 120 + 124 and 123 + 124, respectively. The sequence of the fragments were determined using the ALFexpress AutoRead kit to fluorescently label the DNA, which was analyzed on an ALFexpress sequencing apparatus (Pharmacia). The data indicated both PCR fragments possessed significant homology to polyketide synthase genes of Type I antibiotics. The 650-bp fragment was chosen for hybridization experiments.

25

15

30

35

40

25

45

50

55

The fragment was labeled with <sup>32</sup>P-dCTP using the NEBlot kit (New England Biolabs, Beverly, MA) and purified on a Bio-Spin 6 column (BioRad, Hercules, CA.). Duplicate blots were pre-hybridized in 3 X SSC (1 X SSC contains 0.15 M sodium chloride and 0.015 M sodium citrate, pH 7.0), 4 X Denhardt's solution (100 X is 2% Ficoll [Type 400], 2% polyvinylpyrrolidone, and 2% bovine serum albumin [Fraction V]), and 100 µg/ml sheared, denatured salmon sperm DNA; all reagents purchased from Sigma Chemicals, St. Louis. The labeled DNA was heated in a boiling water bath for 5 min to denature the strands, cooled on ice, and added to the pre-hybridization solution. The filters were incubated for at least 18 hr in a roller bottle hybridization oven. They were transferred to new bottle, then washed two times in 2 X SSC, 0.1% SDS at 70°C for 30 min (moderate stringency). The membranes were placed on Whatman 3MM paper to remove excess liquid, covered with Saran Wrap, and exposed to autoradiography film (Kodak X-OMAT LS) with two intensifying

5

20

10

screens. The cassette was placed at -70°C and developed at appropriate intervals.

15

20

25

30

35

40

45

Approximately 100 colonies were seen to have hybridized on the duplicate filters. Fourteen of these were isolated from the master plate and grown in 4 ml LB + kanamycin medium for 20-24 hr, 37°C, 250 rpm. Plasmid DNA was prepared using the standard alkaline lysis method and digested with restriction endonuclease PstI. The digested DNA was electrophoresed on a 0.8% agarose gel in TAE for 3 hr at 100 V. Fragments were transferred to Duralon UV using the VacuGene XL vacuum blotting unit (Pharmacia) and the recommended alkaline denaturation protocol. Hybridization with radioactively-labeled PCR fragment and washing were carried out as described above. Two prominent types of cosmids were observed; one contained PstI fragments of ca. 7.0, 5.0, and 1.1 kb (pEPOcos6 and pEPOcos7) that hybridized to the probe; the other type had fragments of ca. 6.0 and 3.6 kb (pEPOcos8 and pEPOcos13) which were homologous to the probe. Restriction analysis confirmed that cosmids showing identical hybridization patterns had identical or overlapping inserts. PCR reactions using primers representing consensus sequences of Type I PKS genes were performed using the isolated cosmid DNA as template under conditions described above, except ca. 0.01 µg of cosmid DNA was included as template. Cosmids pEPOcos6 and pEPOcos8 amplified the 650-bp fragment seen when oligonucleotides 120 + 124 were used, while pEPOcos8 and pEPOcos13 supported amplification of an 1100-bp PCR fragment with oligos 122 and 124. The latter fragment was sequenced and confirmed to possess strong similarity to Type I PKS genes. These data confirm that the recombinant cosmids are related to each other and that all contain PKS-like genes.

50

55

5

21

10

## 2. Colony blot hybridizations of plasmid library in pZero2.1:

15

20

25

A 137-mm circle of Duralon UV membrane was placed on top of a 150-mm containing 75 ml LB agar + kanamycin. An aliquot of the plasmid library (representing ca. 2,000 recombinant colonies) in 0.5 ml LB medium was spread on the filter. The plate was incubated at 37°C for 20 hr. Colonies were replicated onto two fresh filters which were placed onto LB + kanamycin agar medium and incubated at 37°C for 6 hr. The filters were processed for hybridization as described in Section C.1. Out of 8 positive colonies detected, one contained a plasmid with a DNA region not encoded by either pEPOcos6 or pEPOcos8. This plasmid, called Sau4, was characterized in more detail.

15

## 3. Colony blot hybridizations of cosmid library in SuperCos:

30

35

40

45

30

50

55

The recombinant *E. coli* clones from the microtiter plates (see section 4. b) were used to produce two identical sets of hybridization filters in order to identify cosmids carrying PKS and PS genes. The recombinant clones were spotted onto 2 sets of 22 x 22 cm LB agar plates containing 50 µg/ml kanamycin. Each plate contained 384 clones therefore representing 4 microtiter plates. The clones were incubated at 30°C overnight. After pre-cooling for approximately 3 h at 4°C, 20 x 20 cm Hybond N<sup>+</sup> Nylon membranes (Amersham, Braunschweig, Germany) were placed onto the agar surfaces. After 2 min. the membranes were removed and placed for 15 min. on Whatman 3 MM paper (Whatman paper Ltd., Maidstone, England) soaked with denaturation solution (0.5 N NaOH, 1.5 M NaCl) before they were transferred onto Whatman 3 MM paper saturated with neutralization solution (1 M Tris-HCl, pH 7.5, 1.5 M NaCl). Subsequently the membranes were

5

22

10

15

placed onto Whatman 3 MM paper soaked with 2 X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.2) for 10 min. The membranes were baked for 40 min at 85°C. Then, each membrane was overlaid with 5 ml Proteinase K solution (2 mg/ml Proteinase K in 2 x SSC) and incubated at 37 °C for 90 min. Finally, cell debris was removed by wiping the membranes with a Kimwipe pre-wetted with 2 X SSC.

20

25

30

As we were seeking in particular to identify biosynthetic pathways containing both PKS and PS genes, the following hybridization strategy was taken: The screening was initially focused on ketosynthase domains from type I PKSs and on the adenylation domain from PSs. Target-specific primers were used to amplify DNA fragments of the corresponding genes from chromosomal DNA of *S. cellulorum* by PCR. The fragments obtained were then cloned, sequenced and the deduced amino acid sequence compared to known ketosynthase and adenylation domains of PKS and PS, respectively. In a second step these PCR fragments were used as gene probes to detect recombinant cosmids of the *S. cellulorum* cosmid library.

35

40

Oligonucleotides based on conserved amino acid sequences of ketosynthase domains from various type I PKS were optimized for myxobacterial DNA by comparison to a known myxobacterial biosynthetic gene cluster (Schupp et al., J. Bacteriol. 177, 3673-3679 [1995]) resulting in primer

25 KS1Up (5'-

C/A)GIGA(A/G)GCI(A/C/T)(A/T)I(C/G)(C/A)IATGGA(C/T)CCICA(A/G)CAI  
(A/C)G-3') and

45

KSD1 (5'-GG(A/G)TCICCA(A/G)I(G/C)(T/A)IGTICGIGTICC(A/G)TG-3').

PCR-primers TGD (5'-

30 T(A/T)(C/T)CGIACIGGIGA(C/T)(C/T)(G/T)IG(G/T)ICG-3') and

50

55

5

23

LGG (5'-

10

A(A/T)IGA(A/G)(G/T)(G/C)ICCI(CI(A/G)(A/G)(G/C)I(A/C)(A/G)AA(A/G)  
)AA-3')

15

directed to genes encoding adenylation modules have been de-  
5 scribed by Turgay et al. (Pept. Res. 7, 238-241 [1994]). PCR  
reaction mixtures with a final volume of 25  $\mu$ l contained 0.1  $\mu$ g  
template DNA, 0.2 U Taq DNA-polymerase (Gibco BRL, Eggenstein,  
Germany), 5  $\mu$ mol dNTP, 5% dimethyl sulfoxide (Sigma), 1.5 mM  
MgCl<sub>2</sub>, 25 pmol of each primer and the appropriate reaction  
20 buffer supplied by Gibco BRL. Chromosomal DNA of *S. cellulosum*  
was used as template. Additionally, chromosomal DNA of *Myxococ-*  
*cus fulvus* was used with PS primers. Reactions were carried out  
in an Eppendorf Mastercycler Gradient (Eppendorf, Germany) using  
the following conditions: denaturation 30 s at 97°C, annealing  
15 30 s at 55°C, extension 60 s at 72°C for a total of 30 cycles.

30

The formation of ca. 700 bp fragments using the KS primers and  
of ca. 350 bp fragments with the PS primers were confirmed by  
0.8% agarose gel electrophoresis. Fragments of independent PCR  
reactions were ligated into vector pCR2.1TOPO using the TOPO TA  
20 Cloning kit (Invitrogen, Leek, The Netherlands) according to  
the manufacturer's protocol and transformed into *E. coli* XL1-  
Blue. Sequencing of the resulting plasmids and analysis of the  
deduced amino acid sequence revealed three different KS frag-  
ments, designated pM008.4, pM008.6, pM008.7, one PS fragment  
40 (pAPs1) corresponding to *S. cellulosum* and one PS fragment  
(pDPs1) obtained with chromosomal DNA of *M. fulvus*. The PCR  
fragments were re-isolated by digestion with EcoRI from the  
45 plasmids pM008.4, pM008.6, and pM008.7, labeled, pooled and  
used as gene probes in hybridization experiments as described

50

55

5

24

below. The same procedure was performed with the PS fragments of pAPs1 and pDPs1.

10

Hybridization with PKS and PS specific DNA probes (see above) was carried out using the DIG nonradioactive labeling and detection kit (Boehringer Mannheim, Germany) and performed according to the supplier's manual using buffer containing 50% formamide. The membranes were hybridized in plastic bags containing approx. 10 ml of hybridization solution at 39°C overnight. Unspecific binding of probes was removed by 2 wash steps with 2 x SSC, 0.1% SDS at room temperature for 20 min. and one stringent wash step with 0.5 x SSC, 0.1% SDS at 60°C for 20 min. Detection of hybridizing DNA fragments was performed with the above mentioned system according to the manufacturer's protocol using CSPD as chemiluminescent substrate. The signals were recorded by exposure of the treated membrane to Hyperfilm ECL (Amersham Life Science, Little Chalfont, England) which was developed in appropriate time intervals.

25

30

35

40

45

71 signals were detected with the PKS specific gene probe. On the duplicate filters 35 signals were obtained with the PS specific gene probe of which 7 were already known from the PKS hybridization experiment. These recombinant cosmids harbored PKS- and PS-encoding genes. In order to corroborate these results PCR experiments were performed with DNA of the 7 recombinant cosmids as template and PKS (KS1Up, KSD1) and PS specific primers (TGD, LGG) generating fragments of the expected size of approx. 700 bp and 350 bp, respectively (primers and reaction conditions see above).

25

30

A comparison of the restriction fragment patterns of the DNA from the 7 recombinant cosmids carrying PKS and PS genes digested by BamHI facilitated an arrangement of the cosmids in

50

55



5

25

10

3 groups. They were represented by cosmids designated A2 and A5. The remaining group was represented by pEPOcos6. Therefore, A2 and A5 represented good candidates for further DNA sequence analysis because they carry both PKS and PS genes.

5

15

#### D. Random "shotgun" sequencing of recombinant cosmids and plasmids

20

##### 1. Library construction

10

25

30

a. pEPOcos6, pEPOcos8, A5, and Sau4: pEPOcos6 and pEPOcos7 were sequenced to completion, and contiguous sequence data and analysis for these overlapping cosmids is presented below for the "cos6 region" (cf. claims 7 and 9). Sequencing of cosmid A5, pEPOcos8 and plasmid Sau4 was taken to the point of large contiguous sequences (contigs) representing the *S. cellulosum* insert; sequence and analysis presented below (cf. claims 10 to 15).

35

40

45

Randomly sheared libraries were constructed for cosmids and plasmids of interest using a protocol similar to that of of Fleischmann et al., 1993 (Science 269, 496) and modified in Fraser et al., 1995 (Science 370, 397). Briefly, Qiagen-column purified cosmid DNA (~10 µg) was sheared to a size of approximately 2 kb and the DNA end-repaired using BAL31 nuclease. The DNA was gel-purified after electrophoresis through a 0.75% low-melting temperature agarose gel containing 0.5 µg/ml ethidium bromide in 1X TAE buffer run at 80 V for 2 hours. The volume of the low-melt agarose gel slice was estimated by adding the gel slice to a microfuge tube and weighing, then 0.1 vol. of 3 M sodium acetate (pH 7) was added and the agarose incubated at 60°C. The temperature was equilibrated to 37°C, and DNA ex-

30

50

55

5

26

10

15

20

25

15

30

35

40

25

tracted twice using an equal volume of buffered phenol (Life Technologies). The aqueous phase was transferred and extracted once with an equal volume of chloroform, then ethanol precipitated by the addition of 2 vol. cold 100% ethanol. DNA was concentrated by spinning at 16,000 x g in a microcentrifuge. The DNA pellet was washed with 1 ml 70% ethanol and resuspended in 100  $\mu$ l of 0.1X TE. The DNA was ligated to SmaI-digested, phosphatase-treated pUC18 vector (Pharmacia), and single insert recombinants isolated by gel-purification of the band containing vector plus a single insert, followed by T4 polymerase polishing, and a final intramolecular ligation of the vector-plus-single-insert DNA. This final ligation represents a library of highly random ca. 2 kb fragments that was used for shotgun sequencing of the ca. 40 kb cosmids or ca. 10 kb plasmids.

b. Cosmid A2: Cosmid DNA with inserts of *S. cellulosum* was isolated by an alkaline lysis procedure and purified with Macherey Nagel columns (Machery und Nagel GmbH und CoKG, Düren, Germany) using manufacturer's recommendation. Purified Cosmid DNA was sonicated, end-repaired using T4 DNA Polymerase (Boehringer Mannheim, Germany). After gel-purification fragments of a size of approximately 2 kb were ligated into SmaI-digested, phosphatase-treated pTZ18R vector (Pharmacia). The ligation represents a library of highly random ca. 2 kb fragments that was used for shotgun sequencing of the ca. 40 kb cosmid.

## 2. Sequencing and assembly

45

a. pEPOcos6, pEPOcos8, Sau4, and A5: DNA (1  $\mu$ l of 100  $\mu$ l total in the library) was transformed into *E. coli* by electroporation (20  $\mu$ l of Electromax DH10B cells from Life Technolo-

50

55

5

27

gies) and cells spread onto LB plates containing 50  $\mu$ g/ml ampicillin. After growth overnight at 37 °C, transformants (ca. 300-3000 CFU total) were transferred to 96-well growth blocks and shaken overnight at 37°C in 1.3 ml LB medium with 50  $\mu$ g/ml ampicillin. Templates were prepared from these cells by an alkaline lysis procedure (Qiagen QiaQuick Turbo Prep) to yield purified, double-stranded plasmid DNA. Cycle-sequencing of the plasmid templates was performed using universal forward and reverse primers and BigDye Terminator sequencing kits (Applied Biosystems), using the manufacturer's recommendations, then resolved using an ABI377 automated sequencer. Sequences were edited using Phred, then assembled into larger contiguous sequences using Phrap (Phil Green, University of Washington, St. Louis, MO).

15        b. Cosmid A2: DNA (1  $\mu$ l of 20  $\mu$ l total in the ligation) was transformed into *E. coli* DH10B by electroporation and cells were spread onto LB agar medium containing 50 mg/ml ampicillin. After growth for 18 hr at 37°C, transformants were transferred to 96-well growth blocks and shaken overnight at 37°C in 1.3 ml 2x YT medium with 50 mg/ml ampicillin. Templates were prepared from these cells by an alkaline lysis procedure (Qiagen Qiaquick Turbo Prep) to yield purified, double-stranded plasmid DNA. Cycle-sequencing of the plasmid templates was performed using universal forward and reverse primers and Big Dye Terminator sequencing kits (PEBiosystems) or Thermo Sequenase fluorescent labelled primer cycle sequencing kit (Amersham Pharmacia Biotech) using the manufacturer's protocols. In the shotgun phase of a cosmid, identical amounts of samples were sequenced either by dye-primer or dye-terminator chemistries (Pharmacia, PE Biosystems). Data were collected using Licor and ABI 377

50

55

5

28

10

automated sequencers and assembled with the GAP4 program (Bonfield, Smith, Staden, Nucl. Acids Res. 23, 4992-4999 [1995]). Gaps were closed using custom made primers (MWG-Biotech) on plasmid templates or PCR products in combination with dye-terminators.

15

#### E. Bioinformatic Methods

20

##### 1. Open reading frame (ORF) identification

10

ORFs were identified in the pEPOcos6 region using the OMIGA 1.1.2 (GCG 0.4D) program from Oxford Molecular Limited. Default values were used (Standard genetic code, all ORFs over 50 bases) to generate ORFs; analysis of these results lead to the list of 14 highest quality ORFs as defined in claim 9. Other ORFs, genes, or genetic elements may be found in the pEPOcos6 insert that have not yet been annotated. In addition to hand-editing of the OMIGA-generated data, the MAGPIE automated genome analysis tool:

25

30

(<http://genomes.rockefeller.edu/magpie/magpie.html>)

35

20

was used to identify genes for all the sequenced cosmids and plasmids. ORFs identified in this manner are presented as both nucleotide and peptide files below.

40

For cosmids A2 and A5, ORFs have been identified within the DNA sequences of A5 (contigs 10, 11, 12) and of A2 using the FramePlot analysis program from Ishikawa and Hotta (FEMS Microbiol. Lett., 174, 251-253 [1999] public available under (<http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl>) which is based on positional base preference in codons typical for organisms having genomes with a high G + C content (Bibb et al., Gene 30,

45

30

50

55

157-166 [1984]). Default parameters using ATG and GTG as start  
codons were used. The deduced amino acid sequence of predicted  
ORFs were compared with protein databases (GenBank, CDS trans-  
lations, PDB, SwissProt, PIR, PRF) using BLASTP (Altschul et  
al., Nucleic Acids Res., 25, 3389-3402 [1997]). Additionally,  
high scoring amino acid sequences were analyzed using the Pfam  
program [<http://www.sanger.ac.uk/Software/Pfam/>], which identi-  
fied specific domain structures of the submitted proteins  
(Bateman et al. Nucleic Acids Res., 27, 260-262 [1999]).

## 2. BLAST searches

BLASTP2 similarity searches were performed using the peptide  
files from the above ORF identification strategy as query se-  
quences. Searches were performed using the in-house Bioinfor-  
matics BLASTP2 (Version: BLASTP 2.0a19MP-WashU) web page at the  
Bristol-Myers Squibb Pharmaceutical Research Institute (allows  
BlastN2, BlastP2, BlastX2, TblastN, and TblastX searches). In  
addition, peptide files generated by the MAGPIE analysis were  
automatically searched using a FASTA algorithm.

## 3. Best match and probable identification

Analysis of the BLASTP2 and FASTA output led to an assign-  
ment of a best match and probable function. The best match was  
usually the top scoring match, although sometimes another match  
was given because it was a more relevant homolog, or no match  
was found with a significance greater than  $>e-4$ . Probable func-  
tion represents the best estimate of function given the initial  
analysis of the BLAST data and the published literature regard-  
ing the best match, and may not necessarily represent the true  
function of the gene product (hypothetical proteins are of un-

5

30

10

known function). A higher probability score indicates a higher likelihood that the probable function corresponds to that of the best match; e.g., the polyketide synthase matches are all above  $e^{-100}$ , and given the very high significance scores are presumed to function as polyketide synthases (as are the high scoring peptide synthetases).

15

The following is a summary of the sequence data from the pEPOcos6 region, pEPOcos8, A5, Sau4 and A2.

20

10 a. Data from pEPOcos6 region:

25

Summary: A large PKS/PS cluster spanning multiple cosmids. An IS element (designated IS-Sc1 here) is found in the cluster - this may be a potential tool for genetic analysis of *Sorangium*.

15

30

Statistics: Sequence was assembled from over 2000 random sequences (forward and reverse reads of the ca. 2 kb cloned fragments derived).

20

35

47,713 nucleotides of contiguous sequence (no pFD666 vector included)

40

DNA sequence data are as defined in claim 7.

25

Note: pEPOcos6\_ORF7 sequences (cf. claim 9): the predicted N-terminus of ORF7 shows 145 nucleotide overlap with ORF6.

45

Note: pEPOcos6\_ORF8 sequences (cf. claim 9): >pEPOcos6\_ORF8.seq ("ORF9\_up" in Fig.2)

30

50

55

5

31

10

67.3% G+C

15

Table 3 shows ORF data summary. Note: pEPOcos6\_ORF1.seq is truncated at its 5' end; correspondingly pEPOcos6\_ORF1.pep is truncated at its N-terminus.

**b. Data from pEPOcos8 region:**

20

Summary: Two PKS genes found on a cosmid. A Tn1000 insertion is also found (occurred during *E. coli* propagation). No peptide synthetase genes were found; one P450 hydroxylase was identified.

25

30

Statistics: 1952 random sequence reads from the pEPOcos8 library were assembled using phrap, with 1024 of the sequences assembling into 57 contigs. 12 of these contigs were chosen (totaling 56,537 bp) which each contained >6 reads and consisted of about 1000 bp or more. The sequences of these 12 contigs and the associated ORFs are given below.

35

DNA sequence data from contigs are as defined in claim 10. Table 4 shows more data.

40

pEPOcos8 protein data are as defined in claim 11, i.e. for selected ORFs (polyketide synthase, peptide synthetases, or ORFs with high similarity to known genes).

45

50

55

5

32

c. Data from cosmid A5 insert:

10

15

Summary: A cluster of PKS and PS genes found on the cosmid. Other genes possibly involved in this secondary metabolite production include a downstream lipxygenase gene highly similar to eukaryotic orthologs.

20

25

Statistics: 880 random sequence reads from the A5 library were assembled using phrap, with 530 of the sequences assembling into 12 contigs. 3 of these contigs were chosen (totaling 41,556 bp) which each contained >100 reads and consisted of about 9000 bp or more. The sequences of these 3 contigs and the associated ORFs are given below.

15

DNA sequence data from contigs are as defined in claim 12. Table 5 shows more data.

30

Protein sequence data from selected A5 ORFs are as defined in claim 13.

20

35

d. Data from plasmid Sau4 insert:

40

25

Summary: Insert contains PKS genes on two large contigs - most similar to the soraphen PKS gene from Sorangium.

45

Statistics: 565 random sequence reads from the Sau4 library were assembled using phrap, with 84 of the sequences assembling into 18 contigs. 2 of these contigs were chosen (totaling 6596 bp) which each contained >10 reads and consisted of

50

55



5

33

10

about 1000 bp or more. The sequences of these 2 contigs and the associated ORFs are given below.

15

DNA sequence data from plasmid Sau4 contigs are as defined in claim 14. Table 6 shows more data.

20

10 e. Data from cosmid A2

Table 7 shows ORF data summary

25

F. Construction of suitable recombinant expression vectors

15

#### 1. Expression in Myxobacteria

30

Heterologous expression of the ORFs shown in Figure 1 is performed by using a derivative of plasmid pSUP102 (Simon, R., Priefer, U., Pühler, A., Methods in Enzymology (1986), vol. 118, pp. 643-659). In this plasmid the gene for chloramphenicol resistance is changed for a cassette comprising the gene for streptomycin resistance and the promoter element of the Tn5 transposon. Short homologous genomic DNA segments from the host organism are ligated with the DNA sequences of Figure 1 and with efficient regulatory elements into, for example, the EcoRI restriction site of the vector. Following amplification of the vectors in *Escherichia coli* the DNA is transferred by electroporation of the host cells or by conjugation with *Escherichia coli* S17-I (Simon, R., Priefer, U., Pühler, A., Biotechnology (1983), vol. 1, pp. 784-791).

40

45

50

55

5

34

10

By means of the tetracycline or streptomycin resistance, respectively, mediated by the vector the host cells are checked for integration of recombinant plasmid DNA into the chromosome by homologous recombination.

5

15

## 2. Expression in Streptomyces cells

Heterologous expression of the ORFs shown in Figure 1 is performed by using bifunctional Streptomyces-Escherichia coli cosmids pKU206 and pOJ466.

20

10

## 3. Expression in Escherichia coli cells

Heterologous expression of the ORFs shown in Figure 1 is performed by using "bacterial artificial chromosomes", cosmids (for example Supercos, Stratagene GmbH, Heidelberg) and T7 expression systems (Stratagene GmbH, Heidelberg; New England Biolabs Schwalbach, FRG). Expression of recombinant enzymes occurs in Escherichia coli cells constitutively expressing phosphopantetheinyl transferase required for the formation of holoenzyme polyketide synthetases and polypeptide synthetases.

25

15

30

35

40

45

50

55

Table 3. pEP0cos6 region gene annotation summary (continued).

GENE	S'	3'	# BP	#AA	MW (DA)	BEST MATCH (SCORE)	PROBABLE IDENTIFICATION
ORF1	1'	731	732 <sup>b</sup>	244	26218	ACDA_BACSV (e-22)	acyl-CoA dehydrogenase
ORF2	731	2275	1545	515	56317	AF047828 (e-94)	peptide synthetase
ORF3	2301	3356	1058	352	39092	U57065 2mar (e-29)	antibiotic resistance gene
ORF4	3412	4050	639	213	24819	(no match)	hypothetical protein
ORF5	4248	12101	7854	2618	283282	PKSK_BACSV (e-190)	polyketide synthase
ORF6	12097	14322	2226	742	80545	283857 ppse (e-111)	polyketide synthase
ORF7	14178 <sup>c</sup>	16451	2274	758	82007	(no match)	hypothetical protein
ORF7.1 <sup>d</sup>	15066	14196	1671	557	61320	Y4HP_RHISN (e-28)	hypothetical protein
ORF7.2 <sup>d</sup>	16507	16154	354	118	13102	Y4HO_RHISN (e-13)	hypothetical protein
ORF7.3 <sup>d</sup>	16815	16507	309	103	11079	Y4RG_RHISN (e-13)	hypothetical protein
ORF8	16830	17630	801	123	29633	PKSL_BACSV (e-2)	polyketide synthase
ORF9	17628	25244	7617	2539	279286	pir1373015 (e-160)	polyketide synthase
ORF10	25235	27877	2643	881	97101	AF047828 (e-111)	peptide synthetase
ORF11	27867	32498	4632	1544	165084	AF091251 (e-167)	polyketide synthase
ORF12	32498	39922	7425	2475	267116	AF040570 (e-223)	polyketide synthase
ORF13	40031	45559	5529	1843	71258	AF091251 (e-119)	polyketide synthase
ORF13.1	45599	46018	420	140	14590	(no match)	hypothetical protein
ORF14	46015	47577	1563	521	55671	PKN1_MYXXA (e-34)	Serine/Threonine-protein kinase

<sup>a</sup> The predicted ORF1 gene and gene product is truncated due to cloning of the DNA into the cosmid vector.

<sup>b</sup> 731 is the last nucleotide of the last amino-acid-encoding codon; 732-735

5

36

10

15

20

25

30

35

40

45

50

55

is TGA (termination codon). Termination codons have been excluded in the present annotation.

\* note: the predicted N-terminus of ORF7 shows 145 nucleotide overlap with ORF6

\* note: ORFs on opposite strand - make up a transposable element from bp 16863-14130 (2733 bp, 11 bp terminal inverted repeat) that is similar to IS1131 from *Agrobacterium tumefaciens* (IS-66 like element, 2773 bp, 4 ORFs, 11 bp inverted repeat)

Table 4. pPPOcos8 assembly analysis summary (continued).

a. pPPOcos8 assemblies

ASSEMBLY	SIZE (BP)
Contig43	1017
Contig44	1246
Contig48	978
Contig49	1969
Contig50	2877
Contig51	2319
Contig52	1883
Contig53	4871
Contig54	7257
Contig55	5021
Contig56	10945
Contig57	16154

b. selected ORFs from Contig 56 and 57 of pPPOcos8

gene	5'	3'	# bp	# aa	best match [score]	probable identification
Contig56_003	3	8675	8673	2890	ERY2_SACER (e-300+)	polyketide synthase
Contig56_027	10784	8682	2103	700	pir I60218 (e-300+)	transposon Tn1000 (E. coli)
Contig57_001	92	1210	1119	372		
Contig57_002	222	7001	6780	2259	AL021899 pks12 (e-267)	polyketide synthase (M. tuberculosis)

5

10

15

20

25

30

35

40

45

50

55

38

Contig57_026	7424	6975	450	149		
Contig57_027	7116	8375	1260	419	CPKK_SACER [e-50]	Cytochrome P450 monooxygenase
Contig57_030	9074	8337	738	245		
Contig57_031	8412	9080	669	222		
Contig57_032	8432	9667	1236	411		
Contig57_033	9724	8600	1125	374		
Contig57_037	10621	9755	867	288		
Contig57_039	11085	10618	468	155		
Contig57_040	10687	11091	405	134		
Contig57_041	11887	10904	984	327		
Contig57_043	11520	12998	1479	492	AF072709 [e-42]	unknown ORF (S. lividans)
Contig57_044	13730	12018	1713	570		
Contig57_047	13093	13797	705	234		
Contig57_048	14064	13138	927	308		
Contig57_049	14371	13535	837	278		
Contig57_051	13900	14394	495	164		
Contig57_052	14036	14440	405	134		
Contig57_053	14678	15715	1038	345		
Contig57_056	15211	15900	690	229		

Table: 5. A5 assembly analysis summary (continued)  
a. pEPOcos8 assemblies

contig	bp	ORF	5'	3'	bp	G+C content pos.3	best match [score]	probable function	identified domains using Pfam
10	9435	1	2861	582	2280	88.6 %	AAC44128 [e-170]	peptide synthetase [Saframycin]	AMP-binding aa 88-512; pp-binding aa 631-697
		2	3525	2902	624	87.0 %	no match	OmpA like	
		3	6420	4195	2226	85.1 %	P38370 [3e-04]	unknown	
		4	7734	7135	600	98.0 %	P39760 [2e-21]		
		5	9434	7950	>1495	96.2 %	no match		
11	18023	1	2526	751	1776	97.1 %	CAB38516 [e-134]	peptide synthetase	AMP-binding aa 118-520;
		2	3588	2501	1088	94.5 %	CAB38516 [e-12]	peptide synthetase	DUF4 (domain of unknown function, found to the carboxy side of pp-binding sites) aa 5-380;
		3	5210	3579	1632	97.8 %	S53999 [e-102]	peptide synthetase [Pyoverdine]	pp-binding aa 1-30; DUF4 aa 51-442; conflict. ca. 20 aa at N-terminus are missing (conserved serine residue) conflict!!
		4	6261	3011	2649	64.4 %!!	CAA11039 [7e-55]	polyketide synthase [Rifamycin]	ketosynthase aa 14-439;
		5	8087	8258	1830	94.8 %	CAB06094 [e-144]	polyketide synthase [Phenolphthalein]	
		6	17320	8090	9231	92.8 %	AAD04757 [e-180]	peptide synthetase [Lichenysin]	AMP-binding aa 371-758; pp-binding aa 840-904; DUF4 aa 937-1318; AMP-binding aa 1438-1825; pp-binding aa 1907-1071; DUF4 aa 1959-2388; AMP-binding aa 2485-2889; pp-binding aa 2872-3004; pp-binding aa 3027-3046;
12	15898	1	1	1545	>1545	97.5 %	AAC68816 [4e-74]	peptide synthetase [FK506]	DUF4 aa 45-476;
		2	2883	1549	1335	94.4 %	P43492 [e-07]	cytochrome P450 enzyme	
		3	4659	2911	1749	92.5 %	NP 001130 [5e-56]	lipoxigenase	
		4	5896	7086	1191	95.2 %	no match		
		5	7094	7822	729	90.1 %	no match		
		6	8974	7843	1032	87.8 %	no match		
		7	12001	11252	750	92.8 %	no match		
		8	13533	12181	1353	93.3 %	no match		
		9	15897	13981	>1917	93.9 %	CAA19149 [e-09]	regulator	

43368

Table 5.

## b. selected ORFs from cosmid A5

gene	5'	3'	# bp	#aa	best match (score)	probable identification
Contig10_001	2861	582	2280	759	U24657 saframycin (e-155)	polyketide synthase
Contig10_006	2880	3323	444	147		
Contig10_007	3876	2902	975	324		
Contig10_008	2953	3957	1005	334		
Contig10_009	4357	3299	1059	352		
Contig10_013	4180	6585	2406	801		
Contig10_016	7033	5600	1434	477		
Contig10_017	7043	6138	906	301		
Contig10_018	6500	7219	720	239		
Contig10_019	6954	7295	342	113		
Contig10_020	6982	7797	816	271		
Contig10_021	7819	7040	780	259		
Contig10_023	7415	8029	615	204		
Contig10_024	7794	9435	1642	548		
Contig10_027	9435	8806	630	209		
Contig11_001	770	417	354	117		
Contig11_002	2526	751	1776	591	AL035640 (e-113)	peptide synthetase
Contig11_004	1033	2787	1755	584		
Contig11_005	3500	1977	1524	507		
Contig11_007	3586	2501	1086	361		
Contig11_008	2507	3814	1308	435		
Contig11_011	5213	3579	1635	544	AF047828 syringomycin (e-86)	peptide synthetase



5

10

15

20

25

30

35

40

45

50

55

Contig11_012	6459	3811	2649	882	AJ223012 (e-52)	rifamycin	polyketide synthase	
Contig11_016	6511	5210	1302	433				
Contig11_018	5551	6735	1185	394				
Contig11_021	9899	6258	3642	1213	283857	ppsE M. tuberc	(e-129) polyketide synthase	
Contig11_026	17329	8090	9240	3079	AF047828	syringomycin	(e-300+) peptide synthetase	
Contig11_048	16733	17365	633	210				
Contig11_049	17397	17723	327	108				
Contig12_001	1	1545	1545	514	AF082100 (e-63)	(Streptomyces)	peptide synthetase	
Contig12_002	1369	1	1368	456				
Contig12_003	3	1655	1653	550				
Contig12_005	2317	1361	957	318				
Contig12_006	2883	1549	1335	444				
Contig12_007	1777	4926	3150	1049				
Contig12_009	4659	2911	1749	582	LOX5	MOUSE (e-54)	ARACHIDONATE 5-LIPOXYGENASE	41
Contig12_011	4523	5065	543	180				
Contig12_012	4638	5231	594	197				
Contig12_013	4942	5520	579	192				
Contig12_014	6056	5541	516	171				
Contig12_015	5765	6373	609	202				
Contig12_016	5896	7086	1191	396				
Contig12_017	7095	5899	1197	398				
Contig12_018	5955	7331	1377	458				
Contig12_020	7549	7010	540	179				
Contig12_021	7094	7822	729	242				
Contig12_022	7995	7318	678	225				
Contig12_023	8408	7716	693	230				
Contig12_024	7916	9550	1635	544				

5

42

10

15

20

25

30

35

10751	9732	1020	339
9795	11132	1338	445
10880	12124	1245	414
12184	11252	933	310
12079	14098	2020	674
14098	12248	1851	616

40

45

Contig12_029
Contig12_030
Contig12_033
Contig12_035
Contig12_036
Contig12_039

50

55

Table 6. Sau4 assembly analysis summary.

a. plasmid Sau4 assemblies

Assembly	size(bp)
Contig17	2581
Contig18	4015

b. selected ORFs from cosmid A5

gene	5'	3'	# bp	kaa	best match (score)	probable identification
Contig17_001	2485	1	2485	629	U24241 Sorangium (e-213)	polyketide synthase
Contig18_002	2	1510	1509	502	U24241 Sorangium (e-105)	polyketide synthase
Contig18_010	1494	4015	2522	841	U24241 Sorangium (e-245)	polyketide synthase

Table: 7. ORF data summary from A2 insert

ORF	5'	3'	bp	G + C content pos.3	best match [score]	probable function	identified domains using Pfam
1	1666	1	>1666	94.4 %	P54744 [5e-37]	regulation [serine/threonine protein kinase]	kinase aa 47-294;
2	1605	3338	1734	90.7 %	no match		
3	6100	3398	2703	96.0 %	CAA15124 [2e-180]	acyl-IRNA synthetase	IRNA-synthase aa 27-694;
4	7111	6374	738	94.7 %	AAC32457 [2e-17]	monooxygenase	oxidoreductase FAD/NAD-binding domain aa 110-227;
5	9590	8433	1158	74.4 %	CAB42045 [9e-67]	aminotransferase	aminotran_1 aa 2-385;
6	11393	9855	1539	85.5 %	AAD21754 [e-113]	L-dopa decarboxylase	pyridoxal deC aa 46-411
7	13656	12712	945	86%	CAD41201 [7e-41]	oxidoreductase	short chain dehydrogenase domain aa 53- 240;
8	15374	19984	4611	67.6 %	CAB06094 [e-180]	polyketide synthase	ketoacyl-synthase aa 13-438; acyl transferase aa 533-854; short chain dehydrogenase domain aa 1159-1357; pp binding aa 1451-1515;
9	20003	27889	7887	88.0 %	AAC80285 [e-180]	polypeptide synthetase	DUF4 aa 17-409; AMP-binding aa 507- 905; pp-binding aa 991-1054; DUF4 aa 1067-1466; DUF4 aa 1544-1844; AMP- binding aa 2041-2439; pp-binding aa 2525-2589;
10	28251	29402	1020	86.6 %	BAA13079 [2e-44]	peptidase	
11	31720	30401	1320	98.9 %	no match		
12	31982	32932	951	95.3 %	CAB09733 [2e-63]	sigma factor	Sigma70 ECF aa 17-83;
13	33128	33613	486	95.7 %	no match		
14	33661	34077	417	94.2 %	CAA19900 [e-37]	regulation	
15	35611	35255	357	80.8 %	CAA19859 [3e-15]	regulation	HTH aa 5 11-109;
16	37856	35730	>2127	96.6 %	BAA17885 [2e-50]	regulation	response reg aa 46-159; signal aa 326- 542; response reg aa 580-703;

Claims

5

10

15

20

25

30

35

40

45

50

55

## Claims

1. DNA sequence, the expression products of which cause an enzymatic biosynthesis, a mutasynthesis or a partial synthesis of polyketide or heteropolyketide compounds or are involved therein.

2. DNA sequence according to claim 1, wherein the polyketide or heteropolyketide compounds are epothilones.

3. DNA sequence according to any of the preceding claims, wherein the DNA is derived from myxobacteria.

4. DNA sequence according to any of the preceding claims, wherein the DNA is derived from Sorangium strains.

5. DNA sequence according to any of the preceding claims, wherein the DNA is derived from Sorangium cellulosum.

6. DNA sequence according to any of the preceding claims, wherein the DNA is selected from the group consisting of:

(a) the following DNA sequence:

Seq ID No 1 (A2 cosmid)

GGATCGCGGCGCCCTCGCGCTGCTCCTCGAGCGTGCGGAGGAACTCCACGCCAGGCGCGACT  
TGCCGAGGCCAGGCGCGCCACCACCACCACCGCGTTGCGGAGGGCTCGTCGACGCAATGGC  
GCCACTCGGTCGCGAGCTGCGAGAGCTCGCGCTCCCGCCCCACGCAGGGCGTCGGCTTGCCGA  
GGAGCCGTGGGACGGCATCCGGCTCCTCCTTCGGGCGCGAAGCCAGCACCCCTCCGGGCCCCCT

5

46

10

15

20

25

30

35

40

45

50

55

GTACCGTCTCGAAGCGGCTCGCGAGCAGGCTGGCCGTCGCGTCGTGAGCCGGATCTCCGGCG  
GCGACAGGCCATCTCGCCCGGCGATGAGCTGCGCGACCCGATCGACCAGCTCGCCGACCGGCA  
GCCTCGCCTCGACCTCGGCCAGCCCTGTGCGGACGGACACGGGCACGCCTCCGAGCGCCGCCC  
GCAGCGCGAGGGCGCAGTGGGCCGCCGTGTGGCGAGATCCGTGGGCGACTCGGCGCCGGACA  
5 GCGCGACGAGCCACCAGCGCGCTTGCAGCCGATCGAGGCGCCCGCCGTGGCGCGCCGCGATGT  
CCCGCAGCGCCTCGGCCCCGCGCGGCCGCTCGTCCTCCGAGAGCGTGGCGCCGGCCTCGGCGC  
CGCCGTCTTCGGCCAGGATGACGCACATCACCTTGCCTCGGCCGTCTGTATCGCCTCGCCCCG  
GCGCGGCCGGCGCCGCGACCGCGCTCGCCCCGATCGAGAGCCCCCTCGCCGGCCACGGCGGCGA  
GCTCCGCGCGGGCGGCGGCCGCTCGCGCGGCCGCTCTCCCGGTTCTTCGCCAGCATCCGCG  
10 CCACCAGGCGCTCGAGCGGCTCCGGGATACCGTCGCGGAGCTCCCCGAGCCGCGGCGGCTCTT  
CCAGGACGACCCGCATCAGGAGCGCGAGCGCGCTGTTGCCGAGGAACGGCGGGCGCCCCGCGA  
GGCACTGGAACAGCACGCACCCGAGCGCGAACACGTCCGCCCGGGCGTCGACCGGCGCGTCGC  
CGCGCACCTGCTCGGGCGCTATGTACCCGGGCGTGCCGAGCACGGCCCCGGGCGACGTGAGGG  
TCGGCGCGAGCCGGAGGTGGCGCGCGATGCCGAAGTCGAGCAGCGTGACGCGCTCGACCGCGC  
15 CGCCACGAGCATCAGGTTGCTCGGCTTGAGGTGCGGGTGAACGACGCCGAGCCAGTGGATCG  
CGCCGAGCGTCGTGGCCACGCGCGCGGCCAGCGCCACGCTCTCGGCCAGCGTGAGCGGCGCCC  
CGGCGAGCCGCTCCTCCAGGGTCACGCCGTCGAGCCACTCCATGGCCAGGTACGGCCGCCCTG  
CGCCGTCACCCCGTGCGCCACGTACTGCACCACGCCGGGCGAGCCGAGCGTCACGAGCGCCT  
CCGCCTCCCGCGCGAACCGGCGCAGGTGCTTGGCGCTCGCGCCCTGCAAGACCTTGAGCGCGA  
20 CCGCCTGCCCCGACACCCGGTTCGCGCGCCCGGTACACGTCCCCCATCCCGCCGGAGACGGCGA  
GCCGCTCGATCTCGAAACGATCCTCGATCACATCCGCTGCGCGCATGGCGGTGCCAATGTACT  
CCGCGCGAGCCTCGGGCCCCCGCGGTAAGTGCGGCCCTGCGCCCGGTTGAACGCCAGCCCCGA  
GCGTGACCGCCTCGCGCTCGGGATCCACGGCCGCCGGATCGGTCCACGCCTCGACGAGCGCCT  
GCGTTGAACAACCCGCCACCGGGCGCACGCAGCCGGCATCGCCGCGCTGGCCACCCGGCGCTG  
25 CCGCCCTTAGGCTCACCTCCGCGATGCCCGCTGGTTCAACACGGCAGGTCCCTGCAACCCGG  
CCGATCACTACATGCTCCCGGCCGAGGAGCGCTTGCCCGCAGTGCCCGATCTGGTTCGATCGCA  
AGGCCTACTTCGTCTTGACGCCCCGCGGAGATCGGCAAGACGACCTCGCTGCGCACGCTCG  
40 CCCAGGATCTCACGGCCGAAGGGCGCTACGTGGCCGTCTCGTCTCGGCGGAGGTGCGGCCCC  
CCTTCTCTGACGATCCCGGCGCGGCCGAGCTCGCGATGCTCGCAGAATGGCGCGGCACCGCCG  
30 GCGCGCAGCTCCCCGCCGATCTGCGGCCGCCACCGTTCCCCGATGCGCCCGCCGGTCAGCGCA

5

47

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

TCGGGGCCGCCCTGCGCGCCTGGGCTCAGGCCGCGCCGCGCCCGCTCGTCGTCTTCCTCGACG  
AGGCCGACGCCCTGCGCGACGCGACGCTCGTCTCCCTATTGCGCCAGATCCGCAGCGGCTATC  
CCGACCGCCCGCGTGACTTCCCGCACGCGCTCGCCCTCGTCGGCCTGCGCGACGTGCGCGACT  
ACAAGGTCGCGTCGGTCGACAGCGGCAGGCTCGGGACGTGAGCCCCCTTCAACATCAAGGTCG  
AGTCGCTCACGCTGCGCAACTTCACCCGCGACGAGGTCGCAACACTCTACGCTCAGCACACGG  
CCGAGACCGGTCAGGTCTTCCGGCCGGACGCCGTGGACCGCGCCTTCGAGCTCACCCAGGGCC  
AGCCGTGGCTCGCCAACGCGCTCGCCCGCCAGCTCGTCGAGGTCCTCGTCAAGGACCGCGCGC  
AACCCATCACGTCTGCGAACGTGATCGCGCCAAGGAAATCCTCATCGAGCGGCAGGACACAC  
ACCTCGACAGCCTGGTGGATCGGCTGCGCGAGCCGCGCATCCGCGCGGTGATCGAGCCGATGC  
TCGCCGGCACCGCGTTGCCGAGCGTGCCCCCGACGACCTTCGTTTCGCGATCGACCTCGGCC  
TCGTGCGCATGACCGCGGAGGGCGCCCTCGACGTGCGCAACCCCATCTACCGCGAGATCATCG  
TCCGCGAGCTCGCGTTCCCGATCCGCGCCTCACTCCCCCAGATCAAGGCCACGTGGCTCACGC  
AGGACGGCCGCTCGACGCGGACCGCCTGCTCGACGCCTTCCTCTCCTTCTGGCGCCAGCACG  
GCGAGCCGCTCCTCGGCGCCGCGCCCTACCATGAGATCGCCCCGCACCTCGTGGTGATGGCCT  
TCCTCCACCGCGTGGTGAACGGCGGTGGCACCGTCGAGCGCGAGTACGCCATCGGCCGGGGCA  
GGATGGATCTCTGCGTTCTTACGCGGGCGAGACGCTCGCGATCGAGCTCAAGGTCTGGCGAG  
ACGGCCGCCCCGATCCCGTCGCCGAGGGGCTCGCCCAGCTCGACGAGTACCTGGCCGGCCTGG  
GCCTCGATCGCGGATGGCTCATCCTCTTCGACCAGCGCTCCGGACAGCCCCCATCGCCGAGC  
GCACGCGCCGCGAGCGCGCGCTCTCCCCCGCCGGCCGCGAGGTCGCCGTCAATTCGCGCCTGAG  
GGAGCTCGCCGCGCGGCGAGCGCCCTCCACGAGGGCCGGGCCACCTCGGACAGCGTCTCTACT  
CCTCCGAGGGCCGCGCGCCCCCGGCCCGCGCCGCGCCGCGCCGGCTCCAGCTCGCAGC  
GCACCACCAGGACCTCGCCATCCGCGAGCTCCGGCCGCTCCACGAGCGCGTGCGCGCCCGCGC  
GCACCGCCGTGAGCACGTCTCCAGCGCCGGCTTCAGCCGCGCCAGCGTCGCGGCGTTTCGCCC  
CGAGCGCGAGGTCGGTCACGACGCGCCCCACGCTCGCGCCGAGCTCGCTCTTGCGCTTGTTGA  
CCGCCGCCATCGCCGCGCCGCCAGATCCAGGAGCCCCGGATCCGAAGGCGCCGCGACCGCCG  
CGAAATCCGCCGCTGAAGGCCACTTCGCCCGGTGGATCGAGGTATCGCCCGTCTCCTCCGCGT  
ACACCCAGCGCCAGACCTCGTCGGTGATGTACGGCAGGACCGGCGCGAACAGCCGCAGCAGCA  
CCGACAGCCCCGAGCCGCGAGCGCCGCCACCGCCGAGCCGCGCGCCGCTCCCCGGCGCCGCCCT  
CGCCGCGCGCCCGCGCCTTCGCGAGCTCCAGGTAGGCGTCGGTGAACCAGCGCCAGAAGAAAT  
CCTCGGTCCGCTCGAGCGCCGCCCGCAACTCGTGCTCGTCAACGAGCGCGTCGCGTCGTCCA



5

48

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

CCACGGCCGACAGCTTGTGCAAGAGCGCCCGGTCGAGCTCCTCGGAGATCGGGTGGACCTCCG  
CCGACTGGCTGAGCACGTA CT TGCTCGCGTTCCAGATCTTCGTGACGAGCCGCTTGCCGATCT  
TCAGCACCTTCTCGTCGAACGCCGTGTCCGTGCCGAGCCGCGCGCTCGCCGACCAGTAGCGGA  
CCGCGTCCGAAGAATACGTGTGAGCAGGTGCATCGGCGTGACGACGTTGCCCTTGCTCTTCG  
ACATCTTCTTGCGATCCGGATCGAGGATCCACCCGGAGATCGCGACGTGGTGCCACGGGACCG  
ACGACTCGTGCAGCATCGCCTTCGCGATCGTGTAGAACGCCACGTCCTGATGATGTCGTGGG  
CCTGCGGCCGAGATCGGCCGGAAGAGCCGCGCGTGCGCGCCGGATCGTCCCCCAGTGAG  
AGCTGATCTGCGGCGTGAGCGAGCTCGTGAACCACGTGTGCAAGACGTGCGACTCGGCGGTGA  
AGCCGCCGGGCTGGTCCCGCTGCGACGCCTCGTACCCGGGCGGCACGTCGACCGTCGGGTGCA  
CCGGGAGCATCTCGCGCGTCGCGAGCAGCGGCCGGCTGTGATCCGGGTGCCCCTCGGCGTCGA  
GCGGATAACCAGACCGGGAACGACGCCGAAATACCGCTGGCGGCTGATGCACCAGTCACCCT  
GGAGCCCCCTCGGTCCAGTTGCGGTACCGGAGGCGCATGAAATCCGGGTGCCACTTGATCTTGT  
CGCCGTATTGAGGAGCTCGGCCTTCTTGTCGGCGAGCCGGACGAACCACTGCCGCGTGGGCA  
CGAACTCGAGCGGCTGGTCGCCCCGCTCGTAGAACTTCACCGCGCGCTCGATCGGCCTCGGCT  
CGCCCCGCAGCGCCGGCCCCCGGCCGGCGCCCGCGTGTCTCCTCGCGGCGGAGCAGCTCGA  
CCACCGCCGCGCGCGCCTGCTTCACCCCCCTGCCCTGGAGCGGCGCATACGCGGCGTTGGCCG  
CGGCCGGGTGCGGGCTCTCCACGCGCCCTCGCCGAACGTCACCGGCAGGACACGGCCGTTCT  
TGCCGAGCATCTGCCGGAGCGGGAGCTTCTGCTCCCGCCACCAGATCACGTGGTTCGCGTCGC  
CGAAGGTACAGACCATCAGGATGCCCCGTGCCCTTCTCGCGATCCACGAGCGGGCTCGGGAAGA  
TCGGCACCGGCGCGCGGAAGATCGGGGTGAGCGCCGTCTTGCCGAAGAGGTGCTGATACCGCG  
GGTCCTCCGGGTGCGCCGTGACGCCGACGCGAGCCGCGAGCAGCTCCGGGCGCGTCTGTCGA  
TGACGAGCTCCTCGGCCGTCCCCCTCCACCGGAACGCGATGTCTGGAACGCGCCCGATTGCG  
GGCGATCCTCGACCTCCGCCTGGGCGACCGCGGTCTGGAAATCGACGTCCCACATCGTCGGCG  
CGAAGACCGAGTAGAGGTGGCCCTTCTCGTGGAGATCCAGGAACGACAGCTGCGCCGTCTGCG  
GGCAGTGATCATCGATGGTGGCGTACTCGTTCCGCCAGTCGACCGAGAGGCCACCCGGCGGA  
AGACCGCCTTGAAGACCTGCTCGTCCTCGCGCGTGACCTTGTGGCAGAGCTCGATGAAGTTGG  
GCCGCGACACGATGCGCGGCGGCTCCTTCTTGATCGTCTCCGGCGCGGCTGCGGCAAGGTCA  
GGCCGCGCTCGTACGGCGTGCGCACGTGGTGCGGACGTGGAAGTAGTTCTGCACGCGCCGCT  
CGGTGGGCAGGCCGTTGTCTGTCGCCAGCCCATCGGGTAGAAGATGTTGAAGCCGCGCATCCGGC  
GCTGCGGACGACGACGTCCGTGTGCGTGTAGCTGAAGACGTGGCCGATGTGCAGCGAGCCCG

5

49

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

AGGCGGTTCGGCGGCGGGGTGTCGACGACGAAGGTCTCCTCGCGGGGGCGCGACGGGTTCGTATC  
GGTACGTCCCGTCGGCCTCCCACAGGTTCGGCCAGGCGCAGCTCGGCGACGGGCGAGTCGAAGT  
GCTTCGGGAGCGTCGCGGGATCGATGGAGCGGAACGTCTTCTTGATCGTCACGTGGTCACCTG  
CAGAACAGACCCCGCAGGAACCGCCCGCGGGGCGGGCATCCTACGTTCGTCCCCCGGGTGCCGC  
TCAAGGCGCGCCGCGCCCGCGCGGCGGCGATCCGCGATCGCATCCGCGCATCCGCCAGAGCCC  
GGCGGCTCCGCCGGCGCGCGCGCGCCGTCCGTGGAGCCGAGAGGAGAGGGCCGGCGCCAGGTC  
GTGGAGGACGCCGGCGGCGCCCGCGGAGATCGCGGAGAGGCGGGCGCATCGATCGCGGCGA  
GGCCGGGGGCTCAGTCGTAGCGCTCGACGTGGACGTGCTTGCGGTGGACGCCGAGCTCGCCGC  
GGGCGAGCTCGCGGACGGACGAGACCATCCGATCCAGGCCGCAGATGAAGACGTGCGGCGCCG  
GATCTCCGCTCTTCTCCGCGAGCTCCCGGTAGAGCTCGGGCACGTGCGCCTGCACGTAGCCGC  
GGCGGCCGGCCACGACGGGCCGCGCGCGAGAGCGTGATCTCGTAGCGGATCCGGTCGGATC  
CGCGCGGAGCGCCTCGAGCTCGTCGCGGTAGATGACGTCTCCTCGAAGCGCGCGCCGAACA  
GGATCCACAGGTGGGGCGCGGCCAGCCCCGCGCGCAGGGAGGCGCGCAGCATGCTCCGGAGCG  
GCGTGATGCCGGTGCCGGTTCGCGACGAACAAGGAGGGCGCGGAATCCCCGGGATCGCGGGTGA  
AGAGCCCGTGCGGGCCGATGGCGCGGAGCGTGCGCGCGGGCTCGAGCCGGTGCAAGTGCTCCG  
AGCCCGCCCCGCCCTGCACGAGCGTGACCGCGAGATCGAAGCGGGGCGAGCCGTGCGGCGCGG  
ATGCGATGGAGTAGGCGCGCTTACCTCGCCGCCCGGGAGCGGGAGGACGAGGTTGACCCACT  
GGCCCCCTCGAACAGAAACGACCTCCCGTCGGCGCGCTCGAACGAGAGCTCGCGCACGAAAG  
GGCTGAGGGGCGGGCGGCGACGAGGCGGGCTTCGAACGGTTCGGCGTGATCATGGTCGGGG  
CCCGGCGGGGCTCGGCTGCGAGGCCGCGCGGGTGGCGAGGTCTTACCGCAGCCTGCGCCCCGG  
CCCAATCGCGATCGCCGCGGGAAGGGCGCCCGGAGGGCGCGCAATCGCGGGAATCACGGGC  
TTCCGCCCCGTGCGCCGCGGAGCGCGCGGCGGGCCCGGCGCCCGCGCTCCGGCGGGGAGCC  
GTCGCGGGCTCTACCGCACGCCCATGCGGCGGCGCTGCGGGATGTTACCGCCGGCCGGGAGC  
GATCCTGGTTGGGGAGCGCGCGCGGCGGGCGGGGATCCCGGTGCGCGGGCTTCTGCGCGGGGA  
GCTGCCCTCGCTGAGCCGGGCGCTGGTTCGGGCGACTTGCGCGAGCCCAGCGCGAGATCGGAGA  
CGGGGAGATGCGCGCGTCGCTGCATAGAATCCTCCATGGAATCGGTCATCAACACATCGGGAA  
GAGCACCAGGCTGAAAGAAACCTTCGAAGAACCGGCTCTCATACACCTCCATTTCATCGTGC  
GACCCCGGATTACGAGACGGATCGAACCCGCGAGGGACGCTGGCTCTCTGGGCCTCTCCCTGCT  
CGCTCGACCGGCGCCCTCTCGACGCAACTCCGCCGTTTCGTGCGGACGGGACGGTCCGCCTCGC  
CGCACGCTCCCCGTGAGACGACTCAGCGTCTCGACGTCAGGAGAGATGACGACTCGGCCCCGT

5

50

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

CGCGCCACGACCCTTCCGGCTCGGTGCTTCGAGCGCGCGGCCAGCGAGCGAGGGGCGATCGCC  
AGGAGATCACGAATCTCCCGGCCATCGGCCTCCAGCGCCTCGGGCTCGTTCGCTCGTCCCCC  
GCTCCGTCCCCGCGCGCGCACGACGCGAGCTCGCGCGGGGAACCGCGGGCCGCTGTCGTGGCT  
GCTGATGCGCGACGATACAGGGGGGACGCCGTGCCTACCTGGGCAACAGGCGCTCATCTTCTA  
CCACGGCGAGCACTACGGTGAGTGCTGCCATGAGTAGGCCCTGAGGGTCCGCGCGACGGAGC  
GTGGTGTCAGCGAGAGATGCGCATGGTGGACGCGGGCTACGCGTCGAGAGGGACACTAGCACT  
CGACCTCGATCCTGCCCAGCACTTTTTGTGCGGGAGGGCTGCCCTCCCGCTGGCCGCTGGCCG  
CTGGCCGCTCGCCGCTGGCCGCTCGCCGCTGGCCGCTCGCCGCTGGCCGCTGGCCA  
TGTGCGACGTGAGCTCGAGCAGCCCCGCGGTGACGGACAGACCCCGGAGTTCATCGAGCCGGT  
GATGCCGAACCCGCCAAGCGAAAAACGATCCGTTCCGCAGGTCTGTCCTATCATGCAAGC  
TGCTCGATGCGCTGACAGGCTTCTTCGAGATCCTCGTCGGTCTTTGCGAAGCAAAACCGCATG  
AAGCGACTCCCTGCGTCCCTTCAAAGAAGGCGTCGCCTGGCACGCCCGCCACCCCGGTCTCG  
TCGAGCAAGTAGATGGCTCGCTCTCGACCTGTCCTCCCGGGTAGGCGAGACACATCCGCCAGC  
ACGTAGTACGTCCCTGCGGCACGCAGGGTGGCAAGCCCGCTTCTCCAGCGCCCGACAGAAC  
CGGTCTCGCTTCCGTTTCATATCCCTGGGCAAGCCCCGTGTAACGAGCGAGGAAGGCCGCGG  
ATCCCGGCAGCGACTCCATGCTGCAGCGGCGTCGGCGCGCAGACATACAGCAGGTCTGCTCATG  
GCTCCAATGGCCTTCGCCCACCTGGCATCGGCCACGCTGTAGCCGATCCGCCATCCTGTGATG  
CTGAAGGTCTTCGAGTAGCCGCCTATCGTGATCGTACGCTCGGACATGCGCGGAAGGGAGGCG  
ACGCTGACGTGCTCACGGCCGTCGAAGATAAAGTACTCGTAAATTTCTGTCCTGATCACCATG  
AGGTCATGGTGGCAGGCGAGATCGGCGATCTGTTCCAGCTCCATTCGGCCGAACACCTTCCCG  
GAAGGATTTCCAGGAGAGTTCAACACGATCGCCTTGGTCTTCGGGGTGATCGCGCGCTCCAGC  
TCGTGCGCGTCGACATTCCAGCTCAGGGATCGCGCCGTCACATACCGCGGAACAGCCTCGACG  
GCGAGGATAGCCTGGGCGTGATAGGCATAAAACGGCTCGAAGAGCAGCACTTCGTCCCCAGGA  
TTGAGCAAGGCCATGCAAGTGGCCTGAAAGGCCCTGTGCTCCGGCGCTCACCGTGATGTCA  
GTCTCCGGATCCGCCGCGATGCCATTATGGCGAGCCAGCTTCGCCGCGATCGCATGGCGCAGC  
TCCACGATGCCGTCGAAGCGCGAATATGTATTGCACCCCCGATCCATCGCCTCCTTCACCGCT  
TGAAGGATCACCGAAGGAACTGGGGTATCACAGACGCCCTGGGACATATTGATCCCATGGACC  
TTGGCGCACGCCAGGGTCATGGTACGGATATCGGACTGGGCGAGGCGAGCCGCACGATCACTC  
GGTAGACTCTTCATCAGCGTGCTCCTGCTTCTGTTCTGCGGCTCTGCATGGTGTCTTCGGGTG  
GGCTTGTCAGCTCGACGCGCCCATGCAGCGGCGCAGCCCTAGCGGCCGCAGGTCTGTCCACAC

5

51

10

15

20

25

30

35

40

45

50

55

TTCTTTGATGAAAGCGAGACATTGGGCTTTCGTGCCCTGTTTGCCCGCAGCCCTCCAGCCCCC  
AGGTACGGGCTTGTCGGCGGGCCAGATCGAGTACTGCTCTTCGCCGTTACACGACCTGGCA  
ACGCGTCTTGCTTTCGTTCGTCCCGATTTCATGATTTTCCTCGCCCTTCGTTCAGCGCTGCGCGAG  
CATGAAACGAATCGCTCATCGGCGCACAGGCGCGCGCCGGCTGCCCGGAGGCACTCCACGCC  
5 TCCCTCACGGCAACCTCATCGCTCCGGATGTTCCCGATGGCGACTCGGATCGTGTACCTGCCG  
TGGAGACGGGTATGGGACAAAAATACCCTGCCCGACTTGTTGACCTCGTCCAGCAGCGCCTCG  
TTGAGGCGATCGAGCTCGCGTTCGATCGACTCTCTCTCTGCTCGTCCGCCGACCGCATGATG  
CAAGCGAGCGCGGAGGGCCTCATCGGAAAGCAGACCGTACTGAACGGCGTCGGCGCGAGGCGC  
TCCCAATCGGGATCGGCGTCCACCCACTGGGCGAGCTGCTGCCCAATCGGAGGTGCTCCCGG  
10 ATCCGGGCGCCAGCCCTTCATGCGCGAAGTAGCGCACGATCATCCAGAGCTTCAGCGCTCGG  
AAGCGCCGACCGAGCTGGATACCCAGTCCATGTAATTCGTGACGTCGCCCTCGGTGCGGAGG  
TATTCGGGCACCAGACTGAACGCGCGCTTCAGTCGGTCCGGCGTCACGCACGTAGAGCACGCTG  
CAATCCATGGGGGTGAACAGCCACTTGTGAGGGTTCACTACCAGCGAGTCCGCCCCCTCGCAG  
25 CCCGCGAGCACGTCCCTGTGCTCGGGGACGATCGCGGCCATCCCCGCGTAGGCCGCGTCCACG  
15 TGAAGCCATAGCCCGTGCTCCCGGCAAACGCTGACGATGGCGGGGATGGGGTCGACGCTCGTC  
GTGGACGTCGTGCCACCGTCGCGCGACGCAGAAAGGTTCGGAGGCCGGCCCCGAGGTCCTCC  
ACGACGGCGGGCGCGCAGCGCCTCGGGGACCATGCGGAAGGCCGGATCCGTGGGGATCTTCCGC  
ACCCCTCCTGCCCCGATGCCGAGGGTGATGGCTGCCTTCTCGATGGATGAGTGCGCCTGCTCC  
GACGCGTAGAGTCGCATGCGCCGCTGTCCCGCCATGCCCCGGAGCCGGATGGTCGGCTCGGCC  
20 GAGTCGCGCGCGGCCGCGATCGCGAACCATGCTGGCGGTTCGACGCGGTGTCCATGATCGCGCCG  
35 TGCAAGCCGGCGTCGAGATCCAGCATCTGACGCAGCCAGGAGAGGACGAGCTCCTCGAGCTCG  
GTGGCCGCGGGCGACGTGCGCCATAGCATCACGTTGACGTTGAGGCACGCCGCGAGCAGCTCG  
CCGAGGATCCAGGACCAGACGCCGTGTTTCGCGAAATACGCGAAGAATCGCGGATGATTCCAG  
TGCCTGATCCCCGGCAGAATGATCTGCTCGAAATCGGTGAGCACGGCGTCCATCGGCTCCGGC  
25 TCGACGGGCGGGGTTCGGGGCCAGCCTGCCCTTCACGTGCGCGGGGCGGATCGCGGGAAAGACG  
GGGTATCGATCCGGGTGGCCGAGGTAATCGGCCGCCCAATCGATGATTCTCATAACGATCCGG  
CGGAACCTCCTCCAGATCCATGTCCCCGAGCCGTTCTTTCCGCGGGTTCGCTCACGTCAACCTCC  
45 TCGCCCTGCCAGGACAGGATCCTCGAGGTCCCCTGGCTCCGGCGGTGGAAAGCGCTCCTTGAA  
CGTGAAGGCCCCACGGGGTTCGGTCCGTAGCGCCGAGGTGCTCGAGCCGATCCTGCCCTCGCG  
30 GACCGACGGGATGTGCCCGGGCCGGGACCCACCACAGCACGAGGTAATGCGGCTCGAGATGCTC

5

52

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GAACCACCGAGCGCGCTGTGCGAGGAACGCGGCATGATCCGCGGTGTAGGTGAAGGCGAACAG  
GTGCTCGATGGAGGTCCATACCGACAGGGTCACGAGGAGCCGCTGGTCCGGGTACGGACGGAT  
GGACACAGAGTTCCCCTCGGCCGTCTGCAGGCGCCACACGAACCCCTCGCTCCGATCGGCCAG  
ATGGTTGATATGGTTCGAGCCCCTGGACGAAGCCCCTCCATGATCGGATCCTCCAGCGGAGCGCG  
AATACATGCGAAGTTGTATTGCGCGATGTGGTGGCGATGCTCCGACATGTCGCTTTCCATCTC  
CAGCTCCCGCTCACCAATCCCAGCGCTGCTCCGGGGAGCTCATCAGGGCAGACGCGACATCGA  
TCCCGAAGCTCCGCGCGCATCCCCTCGACGAAGGCGGCCTGGACCGCTTCGGCGACGGATCGGC  
CTGCCTCCGGCAAGACCTCGGAGACAAAGAAGAACCGCCTCGTGGAAGGGACAATCTTGCCCC  
GCTCCGCCTGGCGCCATACGAAGTGCCTCGTCACCAGTCCCTCCGCGTCGGCATAACCGACCT  
CGCCGGCGCCGACCGCCACGCTGCCGCCTGAGCCGAGCTCCACGAACGCCTCACCGCCTCGCG  
AGATCTCGAGGCGAACGTCCGGGGCAGCCAGATCGCCGAGGTCCCAAGCGCCGACGGGGACGG  
CGAACCGCAGCGACAGGAGGTTGTAAAAATCGACGAATGCGTTGATGTGCGGCAGCTCTCCAC  
CACCGAGGACCCGCTTCGCCAGCGCCTCGATCGAGCTCGGAAATTTCTTGCCAGAGACCCCA  
CTCGCTTCATCGCCTCGCGCCAGGCAGCCACGTGCGGATGCGACTGGGCGTTTTCTGCGCCCC  
AGCTCCGTCGCAGCTCCTCCTCGACCTTCGGAGCTCCTCCAGCACGGCCGGCCGCTCTGCGG  
CGTTGTCCAGGCCTTCCCCGTACCCGGTGACCAAGATCATCCAGGAAACGACTCCCAGATTC  
GCGGATCGACGATGAATGCCATGTGCCTCCTGCCCCCTCGAGAGCGATCGCCTCGATCGACACC  
AGGCTGTGGATGCATGAGCCGGGCGGTGCGGACGCAGGACCCCGCTACTCATGGCTCTTCGTG  
GCCGATGAACAGGTCCTCCACCCGTCGATCGTGCTCGGTGCCCCGATCCGTCCAGTCCCACCC  
GCCGGCGACCGCGATGTTTGACCCGAGACGTACGAGGCGCGGTTCGGAGGCGAGGAACGCCAC  
AGCATCTGCGACCTCGCTGGCGCGCCCCAGGCGGCCCATGGGGACGCGCCGCTCCATCCACTC  
CTTCTGCGCGGGCGGAAGGTATCCGTTGTGATGAGCCCTGGAGACACACAGTTGACCAGGAT  
TCCATGAGGCGCCTCCTCCGTGGCCAGGCTGCGCGTGAGGATGAGCACGCCGGTCTTCGCGAT  
CGAGTACGCCGCCACGTTTCGGCGCGCCGCGGATCGCGTACGTGGGGCTCAACCCGATATTGAT  
GATCCGGCCGCTCTTTTCGCTGGCGCATGCGCGCCACGGCCGCGCGACAGAGGTAATGAACGCT  
GCTCAGGTTGCTGTCCATGACGTTGCGCCATTGTCGTCGGTCATCGCCGCAAGCGGCTTGAA  
GAAGAAGTCGCCCACGTTATTGACGAGGATGTCGATGGGGCCAGCTGCGCCTCGACGCTGGA  
GAAGAGCTCCGCGGCCGCGTTGGGGCGGGTGACGTCGGCCTGCACCACCATGGTTTCGTGCCCC  
GAGCGCGCGGATCTCGGCCCGCGTCTGCTCGGCCGCGATCCTTGTTTGAATGGTAATTGACGGC  
GACGTCCGCGCCTTGCTCCGCGAGGCGCAGCGCGATCGCCTTGCCAATTCCGCGCGAGCTACC

5

53

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

GGTGACCAGGGCGACGCGCCCGGCGAGCTCCAGCGATCGCGCCTGTGGCAGGGCCGGAGCAGC  
CTCCTGGTGGAGCTCGACGTCGACGGGGAGCTCCACGTGGTAGCTCGTCTCTCGCGGAGCCGC  
GCAGTACCTCTCGTAGAACGCCTCGAGGACGGGCTCTTCGCGTCGCATGATGTCCGCGTGGGA  
TTCGGCGCTGCGCCACGGATAGAGCACCAGGATCTCGTCGGGGCGCACGGTGCTCTGAAAGAA  
TCGCGCATGACCGCACCAACCCCGGGTGTTCGTGGGCTCCTGGGCCGAGCAGATCATCCATTTT  
TTGCATGATCTGCGTGGCCTCGCCCTCCATGCCGGGTTTGATGCGCCATCGCTCCATAACGAG  
GATCATGTCTTGGCTCCTGTTCGTTCATCGCCGTTTCGATCTGGGGGGGCTGCCCCGCGCTCTCG  
AGGGCGCGCCCCCTTGTATTGGCCGCGGATGGTCTGGGTAGCGCTCGCGAGCTTTCGCTTGTGG  
GCGGCGTTCAGGCTTGCGCCTTGATTGACGAACCGCTCGCAGACGAATGCGTGCGATTTCATAT  
GCGGTCGCGAGCGCCACAGGTAGAGACGGTGGCTACGAGCGAGCTTGGGCGGAAGCGCCCTC  
AAGGCGGGTCGGATGGATTTGAGGGTGCGAATCAAGCGGGCGTGCTCTTCGAAGAAGAGACCG  
CTGAATCCATCCTGGACGAACGGCGGGGCCATGGTAGGTGCGACGATGTTGTTGTAATCATCA  
TTCGAGAAATCGCCGGTAAACCGGAAAGATGCCGCCGTTGCCACCAGAGCTGCGTGGAGATC  
GCGAGCGCCTCTCGCGCCCTCGGAACGTACGGTTGGCGAGCGCCTCCGACAGCCACCTGTGC  
GCCATGAGCAGCGCCTGGACCAGCACGAAGAACGCGTGATGCCCCAATACCCAGCGGCCCAGC  
GCGCCGTCCGGAGCGCCCGCCTCCGCCGGAGGCGCAGCCAGCTGCGGGATGCCATTCCATGAT  
TTTGGCCTTCGCTTGCCGGAGAACTGGTGGAGGATGTCCTCGATGGACGCACAGAGATGCCCC  
ATTTCCATGGGTTCAGGGAAGTACCTTTTCAGGCTTTCGCGGATCATTCGGTAATATGCGACG  
ATCACCGCTTCGCGAGTACGTCTCGAGCGGGTCTGGGGCGGAGACCCGGTGACATGGAAATAG  
GCGTCGTACTCCGCCTCCCGGTCCGACAGGCTCCCCGGGATGTCCGGATCTGCCGCTTGCGCC  
TGCCATCGCTCGAGGATGGGGACGGCGAGGGGGGCGAGATGCTCGGCGAGCACCGCCAGGTCC  
CCCTCCGGCGCATGCGCCAACAAGACCTCGAAGGCCTCGGCTGCGACCTTGAGAGGTCTCGCCC  
ATTCCGACGGCCTCGATGGCTTGGGGCAGCGGTAGATGGATGGTATATTTAGCCATGATTTGC  
CCGAAGATTGCCGCTGCGTCGACAGATCTTTCGCGAGCCGGAACGCCATTTCCACTGCTCTGG  
CTCTCAATATTGAATTGAGCCCTGGCGACTGCCATAGGCCCAGTCGCTCGACACAGTGACGG  
AGCGGCCCCGATGCTTTCTCCTTTTTTAGTCCTGCACCGAATACTTCTGTTGGGCGCCAAAGAT  
CCCTTGCCGAGACTGTCCGGCGAGATGTCGTGTGCGAAGCGTCCGCACGTCCAGCGGGCCCAT  
GCGTTGCTAGAGCATAAAACGGTTCGATGCCTGGTTCGAGAGGGAGACGCGAGGAGCCTCCCTT  
TGGGACGGATGAGGAATTTTCGTGACCGAAATGTCGGCAGGAACAGCGGCGCAGAAGCGGCGCA  
TCGATGGGGAACCATGGGTTACGAAGACATTGATGATAATGTCGACGCAATCGCAATCGTCGC



5

10

15

20

25

30

35

40

45

50

55

GATGAGCGGCCGCTTCCCCGGCGCGAGAAACGTCGAGGAGCTGTGGCAGAAGCTCCGCGCTGG  
CGTGGAATGCGTCGTCACCTTCACAGAGGCCGAGGCGCTCGCCGCGGGGGTGAGCCGCGAGAT  
GCTCGCGAATCCCAGCTACGTGCGCAGAGGCGCGCCGCTCGACGGCGTGGAGCTCTTCGACGC  
CTCGTTCTTCGGGTTCAGCCCCAGGGAGGCAGAGAGCATGGATCCGCAGCAGCGCATCTTCCT  
5 GGAGGTCGCCTGGGAGGCCCTCGAGCGCGCCGGTTACGACCCCGATGCCCATTCGGGGCCTAT  
CGGCGTCTTCGCGGGCAGCGCCCCGAGCGGCTACCACTCCCTGGCGCAGTCCGACCCGGAGAT  
CCTAGGCGCCCTCGGCCACTACCAA CTGACGCTGAACAACGACAAGGATTATCTCACCACACA  
CGCCTCGTACAAGCTCAATCTGCGGGGCCCCGAGCGTGTGCGTGCAGACGTCTGCTCGACCTC  
GCTCGTGGCCGTGGTCATGGCCTGCCAGAGCCTGCTCAACCACGAGTGCACATGGCGCTCGC  
10 GGGTGGCGTGGGGATCCATGCGCATCAGCGGAGGGGCTATCTGTATCAGGAGAACGGCATCTC  
TTCGCCCCGATGGGCATTGCCGCGCCTTCGATGTGGCCGCCAAGGGCACCGTGGGCGGCAGTGG  
CATAGGCATCGTCGTCCTGAAGCGGCTCGCCGACGCGCTCGCCGACGGCGACCACGTGCACGC  
GGTGATTTCGAGGAGCGGCGATCAACAACGACGGCTCGAGCAAGATCGGTTACACCGCGCCGAG  
25 CGTGCAGGGGCAGGCCGAGGTGATCGGCATGGCCCAGGCGCTCGCCGGCGTGGAGCCGGATGA  
CATCAGCTACATCGAGGCGCACGGCACGGGGACGCCGCTCGGCGATCCCATCGAGATCGCAGC  
CCTCACGCGCGTGTTCGGGGCGAAGACCGCACGAAGGCAGTTCTGCGCCATCGGCTCGCTCAA  
30 GACCAACCTCGGCCACCTCGATGCCGCCGCGGGCGTCGCTCGCTGATCAAAACGGTCATGGC  
CCTCGAGCACCGCGAGCTGCCCCCGAGCCTGCACTTCGAGCGTCCGAATCCGAAGCTCGAGCT  
GGAGAGCAGCCCTTTCTACGTCAACACCCGCCTCACTCCGTGGCACGCGGCACGAGGTCCGCG  
20 CCGCGCTGGCGTCAGCTCGTTCGGCATCGGCGGCACCAACGCGCACGTGGTCCTCGAAGAAGC  
TCCGGCCCCCGCCTCCGAGCGGCCCCCTCGCGGCGTTGGCAGCTCCTCACCTCGCGGCTCGCTC  
CGAGGCCGGGCTCGCGCGGGCCACGGCCGACATGATCGAGCACCTCGATCGCCACTCCGGCAC  
ATCGATCGCCGATGTCACGTACACGAGCCACGTGGGGCGCCGGGCCTGGCCCTTCCGGCGAGC  
40 GGTCGTGGCGAGAGCGCCGCGGATCTCCGCGCCGCGCTCGCGAGCGAGGGCTCGCCGCGCTC  
25 GATCTCGTCATGCCAGGCGGCGAGGGAGAGGCCCGTCGTCTTCCTGTTCCCCGGTCAGGGAGC  
GCAGCACCTCTTCATGGCGCGGGAGCTGTACGAGGTCGAGCCGATCTTCCGGCAGTCCCTCGA  
CCGCTGCGCCGAGCTCCTGCGCGGCCCGCTCGGCCTCGATCTGCGGCAGGTCTCTACCCCGC  
45 CGAGGGGCAGCGCGACGACGCCGAGCAGGAGCTCGGTAGGACCGCGATCGCCAGCCCGCGCT  
GTTGCGCATCGAGCTCTCGCTCGCCAAGCTGTGGATGGCCTGGGGGATCGTCCCCCAGGCGAT  
30 GATCGGCCACAGCGTCGGCGAGTTCCCGCGGCTTGTCTGGCGGGCATCTTCCGCGAAGAGGA

5

55

10

15

20

25

30

35

40

45

50

55

CGCGCTCCGCCTCGTCGCGGAGCGGGGCCCTGATGCAACAGATGCCGCCCCGGCGCGATGCT  
GGCGGTGCCCCCTCGCGGAGCCCCGAGCTCGCCCCCTACCTCAGCGACGACATCTCGCTCGCGGC  
GATCAACGGTCCGGCTCTCTCGGTGGTCGCTGGGCCGATCGAGGCCATCGACGCGCTCGCGGC  
CGAGCTCTTGACCACGGGCTCTCGTGCCGGCGACTCCACACGCGGCACGCCTTCCACTCGAA  
5 GATGATGGCCCCCGTCGTTGACGCCTTTACCCGATGCGTGTCCGCGGTGAGCGCCGCCCCGCC  
GTCAGGCCACTTCCTCTCGACCCTGACGGGCGGCTGGATCTCCCCGAAGCAGCGACCATCCC  
CGCATACTGGGCCCCGGCAGCTCGTGAGCCGGTGCGCTTCGCCCAGGCCGTGAGGCAGCTGCT  
GTCCGAGTCGACGTGGCTCTGGCTCGAGCTGGGTCCGGGCCAGACCCTGAGCCCGCTCGTACG  
GCAGCAGGCCCCGCGCGGATGGCGGCCAGGTGGTCGTCGCTCGCTGCCGCGCGGAAGGACGC  
10 GGGCGCCGACCACCTCGCGGTGATCGAGGCGCTCGGCCGTGTCTGGAGCGCTGGTGGGACGGT  
CGACTGGAAGCGCTTTCACGAGGGCGAGGCGCGGCGGGTGCTGCTACCGACCTACCCCTT  
CGAGCGGCAACGATACTGGGCCTCTCCGCGCCACACGAGCGCTCCGCCGAAGCGATAATCAA  
GCCGCTCCTCGCGAAGAACCCAAACGTGCGCGATTGGTTCTTCCCTCCCTGCCTGGCGGCGCTC  
GGATCCTCCGGTCTCGTTCGACGCGCAGGCGGTGACCACGCGGCGCTCTACGTGGCTCGTCTT  
15 CATCGGGGACGAGGGCCTCGGCGCGGCGCTGGTGGAGGGCCTCGCGCGGCGGGGGCACGAGGT  
CGTCGCGGTGGTCACGGGTGAGAGGTTGAGCAGACGGGCACGCAGCGCTACACGATCGATCC  
CGCCGCGAATGGCGATGTTGCGTCCCTCTTCGCGCGGCTCGAAATCGAAGGGCGCATGCCGGA  
CCGGATCGTCCATGCCTTCTGCACGTGCGCTGCGGACGGCGCGCGCATCGAGCGCGGAGCCGC  
GCTGAGATCGAGCGCAGGCTGGGCTTCGATAGCCTCCTCCTCCTCGCCCAGGTGATCGCCGC  
20 ACAAGGCATCCGAAGCCGCTGATGCTCGGCGTGATACGACCCGGGCGCACTCCGTCATCGG  
AACCAGATCATCGAGCCCCCTGCGCGCTCTGGTGCTCGGCCCTGCGCGTCATCCGCAAGA  
AATACCCCATGTCTCGTGCCGGAACATCGATATCGATCTCCCGGGCGAAGGCGGGCGCGCGGA  
GATCGCGGCGCGCCTGATCGCCGATCTGGAGCGAGAGTCGCCCCGACTCGGTGGTGGCCTACCG  
CGGCGGCCGGCGCTGGGTGAGAGCATAGAGCTCACCAGATGTCGGCCGGCGGTGAGCTGGCGC  
25 CGCCCCGCGCCTCCGCCAGCGCGGGGCGTACCTCATTACCGGCGGCCTGGGGGGCATCGGCCT  
CGTGGCTGCAGAGCTCTTGGCCCCGAGAGGCGCACGCACGGCTGATCCTGGTTGGGCGGACAGG  
CCTGCCAGCGCGGCAGGGGTGGGACGACTGGCTCGCGGCGCACGGCGCGGGCGACGCGACGAG  
45 CCGAAAGATCCTCCGGATCCGCGCGCTCGAGGAGGCCGGCGCCGAGGTGAAGATCGCCGCGGC  
CGACGTCTCCGATTTCAATGCGATGCGGAGCGTCATCGAGGAGGCCCGGACGCGCTTCGGCCG  
30 CATCGAGGCGTCATTCACTCCGCCGGCATCGCGAGTGGAGGCATGATCCAGCTCAGGACGCC



5

56

10

GATGGCGGCTTGGCGCGTGATGGCGCCGAAGGTCGGCGGCACGCTCGTGCTCGATGCGCTCCT  
CCGGGACGAGCGTCCCGACTTCCTCCTGATCTGCTCGTCGTTGGCCTCGCTGGTCGGCGGCGC  
CAGGAGGGATGCCGCGTCATCTCGGTGCAATGGGACACGTGGAGTGACGTCGGGATGGCGGT  
GGACTTCAAGCTCCCGGCCGATCTCCAAGAGGGGGCGCCGCGAGAGCCTGAAGCGGGGCATCAG  
CTCGAGCGAGGGCGCCGAGGTGCTCGGCCGATCTTGAGCGCAGGCATGAGCGGGCCGCTGGC  
GATTTGCACGTGCGATCTACCAGCGTACAAGCAGTCTGTACGACACGCCGATCGCAGCACGA  
GCAAACCTCCCGCCGCCCGGCCGATGCACTCGCGCCCAACGACCACGGGAGCCTATGTGCTCC  
CGAGACCGAGACCGAACGGCGCATCGCCGCGATCTGGCAGGATCTCCTCGGCCTCGAGCAGGT  
AGGCGCAAACGACGATTTCCCTCCAGCTGGGCGGCCATTGCTGTTGGCCACGCAAGTCTGTC  
TCGCGTCCTGCAGACCCTCAAGGTGGGGATCTCGTTGCCGAGTTCTTCGATGCGCCGACGGT  
CGCAGGGCTTTTCGCGCCTGGTCGACGCAGCACGGGCCGAAGGCGCCGGACCCGTGCGCGCCGC  
AATCGGCCGTGTCGAGCGAGACGCCTACCGAATCAAGCCGCCCGCGGCCGAACAGGCCGCCCG  
CACCAAGCCGTAACAAGAAGGGGATCGAGTCATGGAACCCGTGCGCGGCGTGACATGAATCA  
GCCCGCAAAGCAGCAGGAGACCTGCGTCTTCCCGACCTCCTTCGCGCAGCGGCGGCTCTGGTT  
CCTCGACCAGCTCGAGCCGGGGAGCGCCGTCTACAACATGCCCGCCTCCTTCGGACGCGCGG  
GCCGTACGACGTGCACTCGCTCGTGCGCAGCGTGAACGAGATCGTGCGGCGCCACGAGTCGCT  
GCGCACGACCGTCGATGTCATCGATGGCGAACCCGTGCAGGTGATCGCCCCCTCGCTGCGCAT  
CGAGGTGCCCCGTGCTGGACCTGAGCGAGATCGACGAGCCGGAGCGAGAGGCGGAGGCCCGGCG  
GCTCATGGCGGAGGAGAGCCGCCGCCCTTCGATCTCACGCGAGGGCCGCTGCTCCGAGCCAA  
GCTGCTCCGGCTCGGCGAGGCCGATCACGTGCTGATCTTGACGATGCATCATATCGTCTCCGA  
CGGCTGGTCGATGGACGTGCTGTTCAAGGAGCTTTCCACGCTCTACGCCGCCTTCCACGAGGG  
CCGCCCGTCGCCGCTCCCGGAGCTGCCGATTCAATACGCCGACTTCGCGGTGTGGCAGCGGGA  
GCTGCTCCAGGGCGAAGTTCTGGAATCGCACCTCGGGTACTGGAGAGAGCACCTCCGCGGCGC  
CCCCACGCTGCTGGAGCTTCCGATGGACCGGCCCGGCCGCGCAGACGTTCCGGGGGCTC  
CCAGCGCGCGTTCCGACTCCCACTCTCCCTGCAACAGGCGGTGCAGGCGCTCAGCCGGCAGGA  
AGGCGCGACCCCTTCATGACGCTGCTGACGGCGTTTCAGCGTGCTGCTCTCGCGTTATGCGCG  
GCAGAGCGATCTGGTGGTTGGCACGCCCATCGCGAATCGCACCCGAGCAGAGCTGGAGGGGCT  
GATCGGCTTCTTCGTCAACATGCTGGCGCTGCGCATCGACCTCGGGGGCGACCCGAGCTTCCG  
CGAGCTGCTCGGGCGGGTGCGGGAGGTGACGTTGGGCGCCTACGCGCACCAGGACCTGCCCTT

50

55

5

57

10

CGAACGGCTGGTGGAGGAGCTGTCACCAGGGCGGAGCCCCAGCCACAGCCCCTTGTTCCAGGT  
GTCCTTCACGTTGCAGAACACCCCGATGGATGCGACGAACAGAGCAGACATTGCATCGGGTGG  
CGCGCCGCTGGTGGAAATGAAGGCGGCGAAATTCGATCTGATCCTGGAGCTCTCGGAATCGCC  
GCAAGGGTTGCTCGGCACGTTGAGTACAACACCGACCTGTTGACGCCGGCACCATCGAGCG  
5 GATGGCCGGCCACCTGGAGGTGCTGCTCTCCAGCGCCGTCGCGGCGCCGGATCGACCCATTGC  
15 GGAGCTGCCGCTCATGGGGGCGGAGGAGCGCAGTCGGGTATTGGTGGAGTGGAATCCACTGC  
CGCGCTGTATCCCGAGGACCATTGCATGCACGAGCTGTTGAGCAGCAAGTGGAGCGGTCGCC  
CGAGGCGACCGCGGTGCTCCTCCAGCAGCAGACGTTGACGTATCGAGAGCTGAACATGCGCGC  
CAATCAGCTCGCGCATCACCTGCGGAGCCTGGGCGTGGGCCCAGAGGTGCGCGTCGGGTGTGA  
20 TCTCGAACGGTCAATCGAGACGGTCGTGGCGATCCTCGGCGTGCTCAAGGCTGGCGGGGCCTA  
10 CGTGCCGCTCGATCCGACGTACCCCGAGCGCGCCTCGGGCTCATGATGGCGGACGCAGCGCC  
CTCGGTGCTGCTCACGCAGGCGTCGCTCCTCTCGAAGCTGCCGCCCCACGGGGATGCAACGCT  
25 GGTACAGCTCGACGCGCTGCACGAAGCGCTCTCCAGGCTGCCACACCATAACCCGCGGAGCGG  
CGTCACCGCCCAGAACCTCGCATACTCATGTACACTTCCGGCTCGACCGGGCGGCCCCAAGGG  
15 CGTGCTCGTCGAGCACCGCGGCCTCTGCAACCTGCCCACCGTGCGAGGCCAAGCTCTATGGAAT  
CGCGCCGGGCGACAGGCTCCTCCAGTTGCGCGCGCTCTGCTTCGACACATCGTTCTGCGAGAT  
30 CGCGCTCGCGTTGCTCTCGGGAGCGACGCTGGTCATGGGCACGGCGGACGAGCTTCTCCCGGG  
ACCTCCGCTGGTTCGAGCTGCTGAAGAAGCACGCGGTCACGGCGATGCTCCTGGCCCCCTACCGT  
GCTCGCAGCGCTGCCAGAACACAGAGCGCGGCGTTGCCGCTGCGCGTGCTCACGATGGCCGG  
20 TGAGGCGTGCCCGGCGGAGCTCGTCAAGCGCTGGAAGGCACCCGACGGCGCCTGTTCAACTC  
35 CTATGGCCCCGACCGAGACGACCATTTGGGCAAGCTCCGACGCGACCTGTCCGACGAACGGAT  
CCCGCCCATCGGCCGTCCGATTGCCAATACGCAAATCTACGTGCTCGACGAAGCGCTCGAGCC  
GGTGCCCATCGGCGTGCCGGGCGAGATCTTCATCGGCGGCGTGGGCGTCGCCCCGGGGATATCA  
40 CGGGCGTCCGGACCTGACGGCCGAGCGATTGTAACCGACCCCTTCGGGCAAACCAAAGGGGC  
25 GCGCCTGTATCGGACCGGCGATCGGGCGCGCTGGCTGCCGACGGAAACCTCGAGTTTCTCGG  
TCGAAACGACGAGCAGGTGAAGGTCCGCGGTGTCGCATCGAGCTGGAGGAGATCCGCGCGGC  
GTTGCTCAAGCACCCGGCGGTGCTCAAGCCGTGGCCGTGGTGGCGGAGGACACGCCGGGGGA  
45 CAAGCGGCTCGTCGCGTATGTCGTGCGACGCGGAGGAGCGCGCTGACCGCCGCGGAGCTGCG  
CCAGTCCGTGAGCGAGCGATTGCCTGCGACCATGGTGCCATCGTCCTTCGTGGCGCTCGACGC  
30 CTTGCCCTGACGCCGAATGGCAAGGTGGACCGCCGCGCGCTGCCGGAGCCCGAGCAGAGCGC

50

55

5

58

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

CGGCGGCGAGGACCACGTCGCGCCGCGCAACGCCGTCGAGGAGGAGCTCGCCAGGATCTGGGC  
GAGCGTCCTCCGGCTCGAAAGGGTCGGCGTCCACGACAACTTCTTCGAGATCGGCGGCGACTC  
GATCCTGAGCATCCAGATCGTGGTGCGCGCGCAGCAGGCAGGGCTGCGCCTCACCCCGCGTCA  
GATGTTCCAGCACCAGACCATCGCCGAGCTTTCGACCGTGGCTAGAGCCGTCGAGGCGGTCCA  
CGTCGAGCAGGACCCGGTGACCGGTCCCGCGCCGCTCACGCCGGTGACGCGCTGGTGGCTGGA  
GCAGGAGGCGGCCGAGCCGCACCACTTCAACCAGTCGATCTTCCTCGAGGTACGCGAGCGGCT  
CGACGAGAGCGCGCTGGAGCAGGCCATCGCGCATCTGATCGACCACCACGACGCGCTCCGGTT  
GCGCCTCGCGCGCGACGAACGCGGCGCCACCAGGTCTTCGCCGCGCCGGGAGGCTCGACCCC  
ATTTACGCGCGTCGACCTCGGGGCGCTGCCAGCGCCGAGCAGATCTCCGCCATGGAGAAGGC  
CGCGAGCGAGGCGCAGGCGAGCCTCGATCTGGCCGCGGGCCCGGTCTCGTCGCGCCGTGCTCTT  
CGACCTCGGCGAGGTGCCCCGCAACGGCTGCTCGTCATCGCCCACCATAATTGCGGTGACAG  
CGTCTCCTGGCGGATCCTGCTCGACGATCTCTTTGGGGCCTATGAGCAGGCGCGCCGCGGCGA  
GGCCGTACGCTGCGGCCCAAGACCACGTCCGTCAAGCGCTGGGCCGAGCTGCTCACCGAGCA  
CGCCGGCTCCGAGGCCGTCAAGGCGGAGCTCGGCTACTGGCTCGACTCATCGCGACGAACGGT  
AGCTCCGCTGCCCCGTGGATCGACGGGCCGCGAGGACGTGTGGGGCTCGGCGCGCCACATCGT  
CGTCTCGCTCACGCCGGAGCAGACGGAGCAGCTCCTGCGCGAGGTGCCGAGGCGTACCGCAC  
ACGGATCGACGACGCGCTCCTCACTGCGTTGCGCGAGGCCATCGCTCGGTGGACGGGCTCGCC  
GGCGGTGCTCCTCGACCTCGAGGGTCACGGGCGCGAGGAGCTCGCCGGCGTAGACCTCACGCG  
CACGGTCGGCTGGTTTACGGCCATGTACCCGATCCTACTCCGCGTCGACGCGGCGGATCCGGG  
TGAGGCGCTCAAATCGATCAAGGAGCAGCTCCGCGCCGTGCCAGGCCGCGGGCTCGGCTACGG  
CTTGTTGCGTTACCTTCGGTCCGATAACCATCGCCGAGGTCCGCGCGTTGCCGAGGCCGAGCT  
CTGCTTCAACTACCTCGGCCAGCTCGATCAGGCGATCCCCGAGGCTGCACCGTTCCGGCCGGC  
GCGCGAGTATCAAGGCTCGGAGCGCAGCCCCGGCGCCCATCGCGCCACCTCATCGAGGTGAA  
CGCGAGCATCGCCAATGGGCGCCTGTACGCCACGTGGACGTACAGCGAGCGCCGCCACGAGCC  
CGAAACCATCGAGCGCGTCGCGGCGAGCTTCGTACGGCGCTCCGCGCGCTCATCGCGCACTG  
CACCTTGCCCGAGGTGCGGCGCAACACGCCTTCCGACTTCGACAAGGTGCGCCTGCGCCAGGA  
GACCATCGATGCTCTCGACGCAATCGACGCGGGCCCCGGGCCGTCTGCGAGGGGGAGCCGAAT  
CGAAGACGTCTACCCGCTCTCGCCGCTCCAGGAGGGCATCCTGTTCCACACGCTCTACGCCAC  
CGATTACAGGCGTATGTCGAGCAGTTCCACTGGACGCTGGAGGGCGATTTTCGACGCCGAGGC  
GTTACCCGCGCCCTCCAGGACGTCTCGCTCGGCATGCCGCCCTGCGCACGTCTGCTCGCCTG

5

59

10

15

20

25

30

35

40

45

50

55

GGAGCGCCTCGATGCTCCACTTCAGATCGTCCGCACGGGCGCGGTCTCTCCCGTCGAGCACCA  
GGACCTACGCGGCCTCGCCGCGGAGGAGCAGACCGCGCACATCTCCCGTTACGTCGAGGCAGA  
GCGCCAGCGCCGGTTCGATCTGCGAAAGGCGCCCTCATGCGCGCCGGGCTGCTCCGGCTCCG  
CAAGGACGCCTGGTGCCTCGTCGAGACCATCCACCACCTGATCCTGGACGGCTGGTCGACACA  
5 AATCTTGCTCAAAGAAGTGTTACGCTCTACGAGGCGCACCGCGGACACCGTGGGCATCTCGC  
GCTGGAGCTCGAGCAGCCGCGGCCCTACGGCGATTACATCGGCTGGCTCGCGAAGCAGGACCA  
GGTGCGCACCGCGGCCTTCTGGCGGCGCGAGCTCGAGGGCTTCTCCGCGCCGACGCCGCTCGG  
CGTCGACCGCGCTGTGCCGCACGACGACGGCGGCCCGCGGTTTGGTTGGCGCCGCATCGCCCT  
CTCGGGCGACGACGCGGCCCGGCTCGCCGCCTTCGCGCGTCAGCATCAGCTCAGATGAGCAC  
10 GCTGGTGCAAGGCGCGTGCGGCTGCTCTTGTCACGCTACAGCGGCGATCCCGACGTGCTCTT  
CGGTATGACCGTCTCGGGCCGCTCGGCGCCGATTCCCGGTATCGAGCGCATGACCGGCCTCTT  
CATCAACACCATTCGGGTGCGCGTGCGCGAGCCTGCCGACCGCTCGGTGCTCGCGTGGCTCAA  
GGCGCTCCAGGAGCAGGAGGAGGCTGCTCGAGCACGAGCACAGCCCGCTGGTCGAGGTCCA  
GGCCCATAGCGACGTGCCGCGCGGGACCCCGCTCTTCGAGAGCCTCGTCGTGTTTCGAGAACTA  
15 CCCGGTGCAAGTTCATCTTCGAGGCCCTCCGGTCGAGGGGCGGACGCGCGCGGAGGAGGGCCT  
CCGCATGATCGATGCGCAGTATATCAGTGATCCACCGTATCCGCTGACGGTCGTGCGGCGCCTT  
CCATGGGACGCTTTATCTCAATATTGGCTACGAGCGCCGCGGTTTCGACGACCAGGCCGTGCA  
ACGGATGATCGGGCACGTACGACGCTGCTCCGGGGCTTCGTGCAGAGGCCCGAGACGTGCGT  
CCGCGATCTGCCGTGCTGACGGCCGAGGAGGAGCGCACCCAGCTCCACGCGTGGAATGCCAC  
20 GCCGCGCCGTATCCCGAGGGCCATTGCATGCACGAGCTGTTTCGAGCAGCAAGTGGAGCGGTC  
GCCCGAGGCGACCGCGGTGCTCCTCCAGCAGCAGACGTTGACGTATCGAGAGCTGAACATACG  
CGCCAATCAGCTCGCGCATCACCTGCGGAGCCTCGGCGTGGGCCAGAAAGTGCAGCTGGGCTT  
GTGTCTCGAACGGTCGATCGAGACGGTCGTGGCGATCCTCGGCGTGCTCAAGGCAGGCGGGGT  
CTACGTGCCGCTCGACCCGACGTACCCAGCGAGCGCCTCGGGCTCATGATGGAGGACGCGGC  
25 GCCCTCGGTGCTGCTCACGCAGACGTGCTCCTCTCGAAGCTGCCGCCCCACGGGGATGCAAC  
GCTGGTACAGCTCGACGCGCTGCACGAAGCGCTCTCCAGGCTGCCACACCATAACCCGCGGAG  
CGGCGTCACGGCCCAGAACCTCGCATAAGTCACTTCCGGCTCGACCGGGCGGCCCAA  
GGGCGTGCTCGTCGAGCACCGCGGCCTGTGCAATCTGCCACCGTGCAGGCCAAGCTCTATGC  
AATCGCGCCGAGCGACCGGCTCCTCCAGTTCGCGCCGCTCTGCTTCGACACATCGTTCTGCGA  
30 GATCGCGCTCGCGTTGCTCTCGGGAGCGACGCTGGTGATGGGCACGGCGGACGAGCTCCTCCC

5

60

10

15

20

25

30

35

40

45

50

55

GGGACCTCCGCTGGTTCGAGCTGCTGAAAAAGCACGCGGTACGGCGATGCTCCTGGCCCCCTTC  
GGTGCTCGCAGCGCTGCCAGAACACAGAGCGCGGCGTTGCCGCTGCGCGTGCTCGCGATGGC  
CGGCGAGGCGTGCCCGGCGGAGCTCGTCAAGCGCTGGAAGGCACCCGGACGGCGCCTGTTCAA  
CTCCTATGGCCCCGACCGAGACCACCATTTGGGCAAGCTCCGCAGCGGACCTGTCCGACGAACG  
5 GATCCCGCCCCATCGGCGGTCCGATTGCCAATACGCAAATCTACGTGCTCGACGAAGCGCTCGA  
GCCGGTGCCCATCGGCGTGCCGGGCGAGATCTTCATCGGCGGCGTGCGGCGTCCCCGGGGATA  
TCACGGGCGGCGCGGACCTGACGGCCGAGCGATTCTGTACCCGACCCCTTCGGGCAAACCAAAGG  
GGCGCGCCTGTATCGGACCGGCGATCGGGCGCGCTGGCTGCCGGACGGCAACCTCGAGTTTCT  
CGGTGCAAACGACGAGCAGGTGAAGGTCCGCGGTATCCGCATCGAGCTGGAGGAGATCCGCGC  
10 GCGGTTGCTGAAGCACCCGGCGGTGCTCAAGCCGTGGCCGTGGTGCGCGAGGACGCGCCGGG  
GGACAAGCGGCTCGTCGCGTATGTCTCGGACGCGGAGGAGCGCGCCTGACCCGCCGCGGAGCT  
GCGCCAGTCCGTGAGCGAGCGATTGCCCGCGACCATGGTGCCGTCTCTCTTCGTGGCGCTCGA  
CGCCCTGCCCTCACGCCGAACGGCAAGGTGGACCGCCGCGCGCTGCCGGAGCCCGAGCGGAG  
CGCCGGCGGCGAGGACCACGTGCGACCGCGCAACGCCATCGAGGAGGAGCTCACACGAATCTG  
15 GGCCGACGTACTTGGGGCAAAGCGGGTGGTGTGCACGACAATTTCTTCGATCTCGGCGGCCA  
TTCCCTGCTGCTCGTCCGGGTGCATGATCGGCTCGGCCAGCGGTTTCGATCGGCCGCCCTCGAT  
GGTCGACCTCTTCACCTATCCGACCGTGGCGTCCGTCCGCGCGTTCTTGGCGAACGGGCGAA  
CGGCAAGCAATCGCCGAGGGAGGCCCGCGGCGGACGTACGGAGCGCGGCCGGCGCCGCTGGA  
GGCGCGGGCGCGGCGGGCGAAGGCCATCCGTGGCCCGACCTGACCCGGGCACCCTTCCAAGCC  
20 CCGCCGTTCTTCGCACATCCGCCGCCCTCGAGCGCCGCGTCCAGCGCCGCCGTTTCGCCGACGAG  
GAGGCGCGAGACGACGGTCCAAGGCCTTCGTGGGCTCTTTGCCCGCAATCCGGAAGCTGCGC  
GGCAGTTCGTGCCCCCTGCAATGCTGCCATTGTAGAGCTCCTCCGCTCGCCGCGGCCTCTTTT  
CTTGCGGCCCCGTCCGCGATTGACCTCACATCCTGATCCCTTCTTGCGTCTGTCAGAAAGTGAT  
TGACGGCCAGCGCCGCGCTTGAGATCTTCCGGCGCGCGGCGATTTTCATCGCTCCGGCGCGCCG  
25 TGA CTGTACCTGCGAAGGGATTATAATGAAACATAACATTGGGTGGCTTCTACCCGCCGCC  
TCGCGACGCTTGCCCTTCGTCCCGGCCGTGCAGCCCGAATCACGGTGAGGATGCGCCCTCCGTGA  
CGTCAGCAGAGAGCGGCGCGGCGCCGAGCGCTGACTGCGTTCGCGCTCGGGGCGAAGCTCCAGG  
CGGCGCTGGACGGCGCCGCCCGCGCGCAAAAGGCTCCGGGAGCCGCAGCGGCGGTCCAGAGCG  
GGGACTGTGTCTGGCGGGGCGCCACGGGCGTCTCGGACCTGGTTCGCGAGCACGCCGACGAAGC  
30 CTGGAGATCTCTTTCGGATCGGCAGCATACCAAGACCTTCGTCTCTACGCTGATACTCATGC

5

61

10

15

20

25

30

35

40

45

50

55

TCCGGGCAGAAGGCCGGTTGTCGCTCGACGACGCGGTGTCGAAGTATGTGAAGGGCATCCCCG  
CCGGCGACCAGATGACGCTGCGCCAGATCCTCGGTACACGAGCGGGCTCTTCGATTACACGT  
ACAGCCCGGCGCTCGGCCAAATGATCGAGGTGGATCCGACCCGCGCCTTCGCGCCGGCAGAGC  
TCATCGCCCTCGCCACGGCCGAGGCGCCGTATTTTCGCGCCGGGCGCGGGTTTTTCGCTATTCTGA  
5 ACACCAATTACATCGTGGCCGGCCTGGTGGCCGAGGCGGTGTCGGGCGGGACGCTCGCCGGGC  
TGCTCCGCACGCGCATCCTAGACCTGTGGGCCTCGCGCACACGTATCTGGACGGCGCCGAGC  
CGCCGGTCCAAGGGCTCATCCGCGGCTACGGCGACTACGGCGCGGGCTTGGTTCGACATCACCG  
ACCAGCTGTCGCCCACCGAGGCGTGGGCCGCGCGCCCTGGTGTGGAACGTCGATGACCTCA  
ATCGCTTCTTTGCCCTGCTCATCAGCCACGAGCTGCTCTCGTCGGACGAGCTTCAGGACATGA  
10 CCACCTGGACCCCGACGATGTGGCCCCACGAGCCCGGATATGGCCTCGGCCTCATCGAGCGCG  
ATTCTGCGCTCGGCTCCCTCAACGGGCACTGCGGAATCATCTGGGGCTTTCAATCGGCGTCGT  
ACGGGGTGCCCGGCGCGGCGACGCGATCACCGCGCTCATCAACCGGAGCGACGGCGACGCAG  
CGCGGCTCGTCGACGAGCTCGCGAAGGTCGTGAAAGAGCGCTGATCGAGGCGGAATGGGAGCG  
25 CTTCGGCGGGTGGTGATGGCGCCCGGCGCTCAGAACGCGACGCGCAGCCCCGCGCTCAGCGGG  
15 CCTGCGCCGGGCGACGCGGCCACGGCGCCCGGACCGACGAGGAGCCGCGCGACGGCGGGCGCG  
CTCGGCGCGTCTGCTCTCGCCGACCCGCGGCTTGCCGAACACGTAGAGCGGCAGGCCGACGGCG  
30 ACCCGGCCACCCCGCCGAGCGCGGTGGCGATCGCCACCTCGGACGCCTCGGCGCGCGCGGCG  
CTGCTGTCGTCTGCTGGCTCGCGAAGACCAGCACCGCGCCGCTGAGGATGGCGGGCGCCCGCCAGG  
GTCGTGAGGACGAGCCCCGAGATCACCATGACCGGGCTGTTCCACTCCGTCTGCTCCGCTCCTCG  
20 AAGTCGCGGAACGCCGCCCTCGCCGCCGCGAGCTCCAGCTCGATCCGGCGCTGCTCGGCGCGC  
35 CGCTCGTCCGCGGAGCCGATCTCGTGGACGCGGCGGGCCTGGCCCTCCAGCGCCGCGATGCGC  
GCCTCGTGC CGCGGCGGGTCTCCTCCACGTGGCCCTGGCGGGACCCCGCCACGGCCGGC  
GCGACAGAGGGCGCCGACGCCGGGTGAGGCGGGCGCCGCGGGCGGCTCCGCGGCCACGGAA  
40 GGCGCCCGCCCGCGGGAGGCGGGGCGGCTCCGCGGCCACGGAAGGCGCCGCGCCGCGGAA  
25 GGCGCGGGCGACTCCGCGGCCACGGAAGGCGCGGCCGCGCGGGAGGCGGGGCGGCTCCGCT  
GCAACGGAGAGCGCGGGCGCCGCGGCAGCCAGGCCAGCGCCACGCGACGACACGACGGCGC  
GCCGCAACCGCGCGCGGGCGCGGAAGCGGAGGTGGACCTGCTCCATGCGCGCAGCGTCGCCC  
45 CTCGACAGGGCCGGGTCAAGGCGCGGGAGTCCGAGAGCACGAGACCTCCGCGCCGAGGAAAC  
AGGCGCGCCGGCGGGCCCGCGCGGC3GCTCGCCGCTCACCCCTCGCGCGGCCGCGCGGGCGCC  
30 GCCTCCCCTCCCCGGCGGGCCGCGCGTCCGCGGCCACGCGGAGCAGCTCCTGGAAGTGCCGCT



5

62

10

15

20

25

30

35

40

45

50

55

CCACCGGGCCGAGGTCGATGCCGTCCATGAACGACGTGAACGCGAAGTACGGCAGCAGCGTCT  
GCCAGGCGCGCAGCCAGCCGGGAGGTAGCGCGGCCGCTCCAGCACGCCGGCGGCGCGCAGCT  
CCGCCAGCACCGCGATGTCGTCCAGCGCGATCCGGCCGCACACGAGCAGCTCCACCTCGTCGC  
GCAGCCCGCCGCGAAGGACCGCGGCCAGGAAGGCGTAGACCTCGAGCAGCCCGGCCAGGTAGC  
5 AGGCGTCCTTGGTGAACGGCGCGCCGCCCTCGACGAGCCCGCCGCGGCACACGCGCTGGGCGT  
CGAAGTAGGCGTCGCGGCGCTCGGCGCCGCGCTCGCGCAGGTGCCGGTACAGGTCGAGGAAGC  
15 TCGCGCCCTGCTCGGCCATGTCCACGAGCCGCACCCGCTCGGCGAGCCGGGTGAGGCGGCCGA  
TGGAGAGCGAGCGGCTGTAGAGCTCGGCGAAGATGGCGAGCCCTCCTGCGTGCGCGTGGTGC  
GGGGGCCGCCCCGAGCGCAGGAACGCGCACCGCGGCTGCGCCGCGCCGTTGTGCGCGGTGAGCG  
20 CGTGCGTCTCGACCTCGTGGTGCCACAGCCCTCGGCCCTCCACGCCGCGAAGGTCGCCTCCG  
GCCGGATGCGGACCCGGCTCATGCCGCGACACCTTGGCCGTGACGCGCGGGTCGACGGTGA  
TCTCGAGGTCGAGCCGCGGCGCCCGCGGCCACGCGGGCGGCGAGCATGTCCCGAGCGCGC  
25 CGGCGTCGAGCGGCTCCTCCTCGGGATCGCTGGCCTCGTCCCAGCCGTGGACGCGCAGGCGCT  
CGGTGAGGTGCTCGGCGAGGTCGATGTTCTGAGCGAGCCGCCGAAGAACC GCGAGCGCGCGC  
15 CGCCGTAGAGCTCCTGCGACCGCGCGGAGAACGCGCGGCTGCCCGCGGCCCTCGAGCAGCTCCG  
CGGCCTGGATCTGCGCGCGGACGTTGTCCCGCAGCCAGCCGAGCGCCGGCGCGTCCCCGTGCA  
30 TGGCCCCGAGGAGCTCGCGCAGCTCGGCGACGCGCCGCGCGAGGCCGTGCGGATCGACGCGGT  
ACTCGACCTCGGGGAGGCGGTCTCTCGCCGGCGGCGAAGAAGCGCTCCTCCACCTCGCGCGGCC  
AGGCGATGTCTCGAGCAGCTTGAGGGCCTTGCCCTCCGCCAGGCGGCCGCCACCCGATCGA  
20 GCTGCTCCAGCACGGCGCGGTGATGCTCATCGAGCGCAGGATCGCCGAAACCGCGAGACGCC  
35 GGAACCGTCATTCCCTCGACGAGGCGCGGATGTCATGTTCCGTGCGTTTTTGGAGCGCCGTC  
GTCGCGCTCGCCTGCGGGCTCCGGCGATCCAGCGCGGTTGCATGCAGCGAGGGTGTTCGGGG  
CTGGCTCGAGAGCGTCTTTGGCCCCACACCCGAGACACGAATGCTCCGCGCCGAGCGCGGTTG  
40 ACCGTGGACCCCGCCGGAGAGCCGATGATACGGTCCGGCCGATGTGCGAGAGTGTAGCTCAACT  
25 CGAAGAACACCGCGCGGCGCTCACCCGACACTGCTACCGGATGCTGGGTTGCGTGGTTCGACGC  
CGACGACGCCGTCCAGGAGACGATGCTGCGCGCCTGGCGGAGCCTGGATAAGTTTCGACGGGCG  
CTCGTCTGCTGCGCACCTGGCTGTACCGCATCGCGACGAACGTCTGCATCGACCTGCGGGCCGA  
45 CCGCGCGCGCCGGGCGCGCCCCATCGAGGAAGGCCCGGTGCGCACGGTGGACGACGCGCTCGA  
GACGCGCCCGCGCACCCACTGGCTGGAGCCCGTCCCCGACGCGCACGCCCTGCCGGCGGACAT  
30 CGACGCCGCGGAGCGGGCGATGCTCCGCCAGAGCATCCGCCTCGCGTTCGTGCGGCGCTCCA

5

63

10

15

20

25

30

35

40

45

50

55

GCACCTGCCGCCGAAGCAGCGCGCCGCGCTGCTGCTCACGGAGGTGCTCGGCTGGTCCGCCGC  
GGAGGTCCGCCACAGCCTCAACACCTCGGTCCGCCGATCAACAGCGCGCTCCAGCGCGCGCG  
GGCGACGCTCGCGAGCCGCGATCTCGGCGACGCGCGCCCTCGCTGCCGGAGCCGCAGTCCGC  
GCTGCTCGACCGCTACGTCAACGCCTTCGAGCGGTACGACGTGACGCGCTCACGGCGCTGCT  
5 GCACCAGGACGCGACCCCTGTCGATGCCGCCGTTACCCCTGTGGCTCCGCGGCCACGAGTCGAT  
CCGCGCCTGGCTCGTGGGCCCGGAGCGGGCTGCCGCGGGTCCGCGGCTCATCCCGACGGCGGC  
15 GACCGGCTCGCCCGCGTTTCGCGCAGTATCGCCCGCGCCGAGGGCGGCCACCGGGCCTGGGC  
GCTCATCGTCCTCGACGTGCGGGGGACCGCATCGTCAGCATGACGTCCTTCCTCGACACCGA  
GACGCTCTTCCCGCGGTTTCGGCCTGCCGCTCGATCTACCGGCGTAGCCGCGGGCGCCCTGCCT  
20 GCCTCGCCGCGGGTGCCCTGCCTAGCCGCGGGCGCCCGGCTGGCCACGGGCGCCCGGC  
CTGGCCACGGGCGCCCGGCCAGCGACGGGGCGACGATTTTTTTCTGAGCGACCGATGAGTCCT  
GACGGGGCCGGGGTCTACGGGGGTGAATCCAACACGGAGGCACCCATGACCGTGACCATCGC  
CAGCATCGATCATCGTGACCAGGACCTCATGACCGGGCCCCAGGCCAAGGCGCCGGCCCGCGC  
25 GGCGCGCCCGACGCGGCGCCGTCCAGGCGAGCCGTGTGGCGGGCCGCGTCCTGAGCGGGCT  
GGCCACGCTGTTCTGACGTTTCGACGCCGCGGTGAAGGTGCTGAAGCTGTTCCCGCGGAGGC  
15 GTCCACCGCCGAGCTCGGGTTCCCGGCGACCTCGTCCCCACCCTCGGCTACCTCCAGATCGC  
TTGCCTCGTGGCCTACCTGATCCCGCGCACCGCGGTGCTCGGCGCGATCCTGTGGACCGGCTA  
CCTGGGCGGCGCGATCGCGATCCACGTGCGGGTCGAGAACCCGCTCTTCAGCCACACGCTCTT  
CCCCATCTACGTGCGCCGCGTTCTCTGGGCGGGGCTCTGGCTGCGCGACCGCCGCGTGCGCGC  
20 GCTGACCGCGAGCCCGTCGTGCGAGGGCCGATGAGCTTCACGTTTCACGAGAGTCCATCACGG  
35 TAAAAGGAGAAGCGAGCCATGACCACAAAGAACCCCGCAAGCTCTTCGTCAACCTGTCCGTC  
CGCGACCTGAAGCGATCGATGGAGTTCTTCAGCAAGCTCGGGTTCGAGTTCAACCCGCGAGTTC  
ACGGACGAGAAGGCCGCTGCGATGGTCGTGAGCGAGGAGGCCTATGTCATGCTCCTCGTGGAG  
TCGTTCTTCAAGACGTTTCATGAAGAAGGAGATCTGCAGCACGAGCACGCACACGGAAGGGCTC  
40 TTCGCGCTCTCGTGCAGCAGCCGGGCCGAGGTGACGACATGGTGAAGAAGGCGGTGCGGCG  
25 GGCGGGTTCGACGCGATGGATCCGCGAGGATCACGGCTTCATGTACGGGTGGAGCTTCTACGAC  
GTGGATGGCCACCACTGGGAGGTGATGTGGATGGATCCCAAGGCGATCCAGCCGTAGCCGACG  
45 GGGCTGGGCGCGCCGCTGGAAGAGCCCCCGTGAGGCGGGGAGGCGGGAGGATCACCGTCTTC  
GTAGCCACAGCGATGCAGTATCCGTGCGGCTTCGTATCGAAGCACGGCTGTTACGGGCGCGT  
30 CAGAGCGCGTCGCGAGGTGATGCCGAGCCGCGAGCGACACGGGCACGAGCGTGGCTCCGATG



5

64

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GAGATGAGCCGAGTCTCGCCCATGGTCTCGGGGTCATGAATGGATGAGTAGGGGACTCGCTCC  
TTCGTACGTCGTGCTCGACGGCGACGGCGAGGCCGAGCTCGAAGTGCACGGGGCCTGGACCG  
AAGATCCAGCTCGCCCCGGCGCGAGCCCCGACGAAAAGCGTGTGCCGTCGACGCCAGGGCCG  
TCGTCCCAGCCGGGCGATCCCACCGCGGTGTAGGTGTGTTTCCCGAAGGAACCCGCGAGCGAG  
AGTCGAAGTCCGACCGGCGCTCGCCACGCGACGCCCCGCTGTGCGCGCCGACGCCGCCGAAGCTC  
TCCCCGAAAGGCTTATCCCCTGTCTCGATGAAGCCACCCACCTCGATGACGCTGATGCGGTAC  
GTGAGCGCGAGATTGAGGTGCACCCCGAGCGCTGTCCGAGCCCCGAGTAGAGGCCGGCGCCCACC  
TGCACGCTGAAATCCATGCTCGGCGCGGATCCGCGCGCAGGAGCGACGCCAGGGGCGCTGCCC  
TCCTGCGCGCGGGCCGTCGCCGACGCAAAGAAAGAGGGCTGTGCGGAAGAATCCAAGCGAGATC  
GATCGAAGTGAGCGCATGTGCGGGCCCTGGAGCATCCGCTGTACCAGGTGCGTCGTATTTCATGC  
GGCGCGCCGCGGGCGCCGCGCGCTGGCCTGTCCGACGCGAGATCACGAATCCGCCATCGCT  
CCCCGCGCGCCGCGGGCGCTCTGGTTTCGCTGCGGGCGTGCGCCGGCGCTCGTGTGGCCCATG  
GCAACCTTGTCGCGGTGTCGCTCGAACAGCACAGAGAGTATCGCGTCCGCAACAACCGCGCGA  
CCCGGCGAGACGCTCGTGGGGCCCCCTGCCTCCCCACTTCATCATAACGCCATCAGGAGCACT  
CGACATTTTCATTTCTTCACCTCCACTGGCTGAGGGCGACGGTGCTCGTCATCGGCCGGTTGCT  
CTGGCGGTTGCTCTGGCGGGGTTTCTGACGCCCGGAACCTAACGCTTCGAGCGCTCCCCCTTGC  
TCTCCCGTTCTTCAGCTCCTCCAGCAGGTGCTCGAGGCGCTCGTAGCTGCCTTCCCAGAAGC  
GGCGGTAGTTGTGAGCCAGCCGCTGGCGTCCTCGAGCGGCTTGGCCTCGATCCGACAAGGCC  
TCCGCTGCGCGTCGCGGCCGCGCGAGATCAGGCCCGCTCGCTCCAGCACCTTGAGGTGCTTGG  
AGATCGCGGGCTGGCTCATCGGAACGGCTTCGCCAGCTCGGTACCGACGCCTCGCCGGACG  
CGAGGCGCGCGAGGATCGCTCGCCGTGTGCGATCGGCGAGCGCAGCGAACGTTGCGTCGAGGC  
GCTCGGACGGGGTCATTGCATAACTCCTTGGTATAAAAACCAAGTTAGTTATACAACCTGGGGC  
CCGGGCGGTCAAGCCTCCAGGCGATGGCGGTTCCGGCCGGGGGCTCCGCTCGCGGCACGCGCG  
CCGCGCGGCTACGTGCGCGGCGCGGTGAGCACGTCTGTCAGCGTGGCGCCGACCACGGGCTTG  
GTCAGGTGCAGGTGGAAGCCGGCCCGCCTGGACCTGGCCTGATCGTCGGGCCCCGCGTAGCCC  
GAGAGCGCCACCAGGTAGAGCGCTTCGCCGCCGGGCGCGGCCCGCGCCCGGCGCGGACCTCA  
TAACCGTCGATGCCGGGCAAGCCGATGTCCACGAAGGCCACCTCGGGGCGCAGCTCCAGAAGC  
TTCTTCACGCCCTCCAGCCCGTCCACCGCCACCGTCACCTCGTGCCCCAGCGCCTCGATGTAC  
GCCCCGATCACCCGGCGCACGTCTCCGCGTCTCCACGACGAGCACCCGGCGCCGGTCAGCC  
GCCGCTCGGGCGCCTCGGCGCGCTGCGCCGGAGGCGGCGGGCTCGTCGCGCTGCGCCGGA

5

65

10

15

20

25

30

35

40

45

GGCGGCCCCCTCGCGCGGCGGGGGCGCCCCGGCGCTCGGGGCAGGCTGCGGCGCCGCCCCGGGG  
CCGAGCGGCAGGCGCACGGTGAACCTCGCTGCCCTGGCCCCGGCCCGGCGCTCGCCGCGGCCACG  
CTGCCGCGCGTGCAGTTCCAGGAGCCGCCGCACCAGCGTGAGCCCGAGCCCCAGCCCGCCCCGTG  
CTCCGGTCGATGGTCTGGTCGACCTGCGTGAACAGATCGAACACCTTCTCGAGCATCGCCGCC  
5 GGGATGCCGCGGCCCCGTGTGCGGCACCCGCAGCACGGCCTCGGGCGCGCCGACCGCCGCTCG  
CGCGTGAGGCGCACCGAGATCGAGCCCCCGGCGGGGTGTACTTTCGCGGCGTTGGTCAGGAGG  
15 TTCGTCACCACCTGCTCCAGCCGCGTCGCGTCGGCCCGCATGCCGAAGTCCCCGGGCCCCACC  
GACAGCGACACGTTCATGGCGCCGGGCTCTGACGGCCGGCCTCACCGCGGCGGCGGCGCTCTGC  
ACCAACCGCCGCGAGATCGACGTCCTCGAGGCGCAGCTCCACCGTGCCCCGCGTGATGCGCGAC  
20 ACGTCGAGCAGATCGTCGACCAGCCGCACGAGGTGGCCCATCTGCCGCCGCGCGATCTCCCGG  
TAGCGCGCCGACGCGGGCCCGTCCGCGTCCGCGTCGTCGAGCAGCGTCAGCGACAGGCTGATC  
GAGGCCATCGGGTTCCGGAGCTCGTGGCGGAGCATCGCGAGGAACTCGTCCTTGCGCTGATCG  
25 GCGAGCTTCAGCGCCTCGACGAGCGCCTCCACGCGCCTCCGGGCGCGCACCTGGTTCGGTCACG  
TCGAACGCGAACACGAAGACGCCCTCGACCGCCCCGTCGCGATCGCGCATCGGCTGGTAGACG  
15 AAGTTGAAGAACACCTCCTCCGTGCTGCCGTCGCCCCGGCGATCGAGCCGCACCGGGAGCTCC  
TTGCCGACGATGGGCTCGCCGGTGCGGACCACCGCGTCGAGGAGCTCCCAGATGCCCTGTCCC  
30 TCGAGCTCGGGGAGGGCGGCCCGGATGGGCTCGCCACGAGCGATCGACCGCCGACGAGCCGC  
TGGTAGAGCGGGTTGACCACCTCGAAGACGTGCTCCGGCCCCGCGGAGGATGGCGATGGGCCCC  
GGGGCCTGTCATGAAGAGGTCGTTTCAGGTACTGGCGCTGCCCCCTCGGCCTCGCGCCGGCGGCGC  
35 GCGAGCTCGACGTGGATGCGGACC CGCGGAGGAGCTCCTTCGCGGAGAACGGCTTCACGAGG  
AAGTCGTCGGCGCCGGCCTCGAGGCTGTCGACGCGCGCCTCCTCGCCCGCGCGCGGAGAGC  
ATCACACGGCGACGCCCGGGTGCGATCGTCGGCGCGCAGCGCCCTGAGCAGGCCGAAGCCG  
TCGAGCCGCGGCATCATCACGTGCTGAGCACGAGATCCGGCGGGTGGGCGCGGGCGCGCTCC  
40 AGGGCGGCCCCGACCGTCGGCCACGGCCTCCACCGTCCACCCCTCCGCCACGAGCAGCCGCAGC  
25 GCGTACTCGCGCATGTCCGCGTTGTCGTCGGCGACGAGGACGCGCCCCGGCAGCCTCCCGGCC  
GGCCCCCTCGCCCGCCGGCCGGGACCCCGGCGCCTGCTCGCCGCGGAGCCACTGCGCGGCCTCG  
TCGAGGAAGGGCGCGGCGTCCCGCCCCCGCGGCCGGCGCCGAGGCCGGCGCGAC

or its complementary strand,

30

50

55

5

66

10

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

15

5 (c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

20

10 (d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

25

15 7. DNA sequence according to claims 1 to 5, wherein the DNA is selected from the group consisting of

30

(a) the following DNA Sequence:

Seq ID No 2 (>pEP0cos6 region)

35

20 GGATCACCTGCGGCGGATCGCCGACCTCGTGCTGGTGTTCGGCTCGCTGGATGAGAAGCCGG  
CGGCGCTACTGATAGAGACGGCGACGCCCGGGCTGCGGGTGGAGCGGTTCGGGAGATGCTCG  
GCTTTCGGGCGGCCCCACCTGGCGAAGCTGTCCTTCGACGGTTGCGAGGTCCCCGAGGCTCAGC  
TGATTGGCCGGCCCCGGCTTTGCGCTGATGTATCTGGCCCCCTACGCCCTGGATTTCGGTCGGG  
TCAGCGTCGCCTGGGCCTGCCTGGGCATGATCCGCGCTTGCCCTGGAGACCTGCGCACAGCACA  
40 25 TCCTCACCCGCCGCACCTTCGGCCACCTGCTAGCCGATCACGGCATGATCCAAACCCTGATCA  
CCAACCTGGGGATTACCAACAGGCGACGCTGCTCCACACGCTGCAGGCCTGCCGCGCCAGGG  
ATCGCGGCGACGTGACCGCCTCCGAGGCCACCCTCGCCGCCAAATACCTCGCGTCGCGGACGG  
CGGTCCAGGAGACGACCAACGCGGTCCAGATCATGGGCGCGCTGGGCTGCGACGAGGAGGGCG  
45 CGATCGCCCCGCACTTCCGCGACGCCAAGACGACCGAAATCATCGAAGGCAGCAACCAGATCA  
30 TCGAGGCGCTGCTGGCCAAGAACATCGCCCCGCGCCGGTCGCGACAACTATCGCCGCTTCCTCG

50

55

5

67

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

ATGCGGAAGTCGAGCCCGGTGCGGGCCGGAGGCGCACCATGACGAGCGCGGTCCCGACGCGTCA  
AACCAGCCTGCTCGACGACTTCGAGCGCGTCCCGACGTCGATCCAGAGCGGATCGCCGTCCA  
CGCGAGCGAGACGAGCCTGCGCTATGGCGACATGAATGCGCGCGCCAACCGCATTGCCACGG  
GCTACGGGCGCGCGGGATCGGGCCCAATCAAATCGTGCGGTGGCGATGGCCCGCACGCCCCGA  
GCTGATGATCGTGCTGTACGGCATCCTCAAGGCCGGCGCGGCTACATGCCCATCGCCCGCGA  
CGCGCCCGCGCTGCGCCGCGATCATATGCTGCGCGAGAGCCAGGCTGCTCTGATGATCGCCGA  
CGAAGAGATCGCGGGACTCGCGGCCCGGGTGCTGACGCGCGCGACCCGTTCTTCGCGGCCAT  
GCCGGACCACAACCCCGAGCCGCGTCACGACCCGACCGACCTGATTTACGTCATCTACACCTC  
GGGCTCGACCGGCCAGCCCAAGGGCGTGCCATGGAGCACC CGCGCCGTGTGGAATCGCCTGAC  
TTGGATGCAGGCCAGTATCCAATCGACACGCAGGACGTGATCCTCCAAAAGACGCCGATCGT  
CTTCGACGTGTCGGTCTGGGAGCTGTTCTGGTGGCCGCTGGCCGGCGCCTCGGTGGCCCTGCT  
GCCGCAATCCATGGAGAAGTTCCCTGGGCGATATCGGCGACGGTGGCGCGGTGCGGGGTGAC  
GGTGATGCATTTTCGTACCATCGATGCTGATGGCCTTCCTTCAGGTGGTGGCGGGCCGGCCCGA  
GATGGCGGACCAGATGAAGGGCCTGCGCTACGTCTTCTGCAGCGGCGAGGCCCTGGCGCCGGC  
CCACGTGTCAGCCTTTCAGGAGCACATCAACCGAGCGGGCAGCATCAGCTTGACCAACCTCTA  
TGGACCCACCGAGGCGGCGGTGACGTCAGCTACTTCGACTGCCCCGCCGGCGCGTCACTCGC  
GCGGGTGCCGATCGGACGAGCGATCACCGGCATCCAGCTGCTGGTCATGCGCGACGGCGTGCC  
TCAGCCGCCCGGCGTTCGAGGGTGAGCTCGCCATCGGCGGCGTTGGTTTGGCGCGCGGCTACAT  
CTCACGGCCAGACCTGACCGCCGACCGGTTCTGTCCGCATCCAGGCGGCGACGGCCAGCGGCT  
CTACCGCACCGGCGATCTGGTGCGCAGGGACGCGGACGGCGAGCTGGTCTTCCTGGGGCGCAT  
CGACCATCAGGTGAAAATTTCGCGTCTGCGCATCGAGCCCCGGGGAAATCGAGGCCAGATCAG  
CGCCCATCCCGATGTGGCCGACTGCGCGCTGATTATCGAGCAGGACTCGGAAACCCTGCCCAA  
GCTGACCGCCTACATTGTCTGTGGCGGACCGGGCTTGACCCGGAAGGCGCTGCTACAGTTCCT  
GGGCGCGCGGCTGCCCGACTACATGCTCCCGAACCGCTTCCTGACCCTCACGGAGCTGCCCCGT  
GACCGCCAACGGTAAGCGCGACTGGCGCGCGCTGCTCGGCCCGCTCGAGACCCTGCCTCTCCC  
TTTCTCCTGAATCCAACCAATACGAGGGATTTCATGTTACACCCGATTCCCACCGACCGTTTCG  
CCCTGAGCCGACCGCTCTTTTCGCGGGTACCTCGCGCACGATCCGATCGTGAGGGCGTGCTGG  
CGGGCGACCATCCAGGCTGGGTCTTGGTGACCGCGAGCCCGAGCCGCGCACGGCGCTGCTGT  
GGGCCTTTTCCGATCGGCTCTTCTGCGTGGGCGCAGCTGACACGCTGACCCCGCACGCGCTGG  
CCGAGCTGTTCCACGACCGACTGATCCCCAGGCCCGTAAGATCGGGCAGCCGTTTTTCCAGG

5

68

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

TTCAGGGCGAGACGGTCGACACCTGGTCGGACCACCTGCATCAGGTGTCGCCGCACGCGACAG  
TCTCCTTCCGCCAGGCATTCCGCTTCGACCGCGACCTCTTCGAGCGGCTGCCAACCAAGCCGG  
AGCTGGCAGAGGCGCGGCTCGTGCCAATCGACGCGCGGCTGCTGGCCGAACAGGCTGATCTGC  
GCGAGCGGATACTGGCCTCCTGGTCCAGCGAAGCTGCCTTCCATGCGCGCGGTTTCGGCTTCT  
GCTACCGCGTAGGTGACCAGCTGCCGAGCGTGTGCCTGGCATCGCACGTAGGCGGCGGCGCGG  
CCGAGCTGAGCATCAACACCGAGCTCGAAGCGCGCAATCGAGGTATGGCAACGCGGCTGTGCC  
GGCGTTTCATCGCCGAATCGCTGCAGCGCGGCTGACGCCTTGCTGGGGCACCGAGACCTTTC  
GCCTGCCGTCAATCGCGCTGGCCCAGAAGCTCGGTTTCATCCCGACCTTCACCTTCCCCACCT  
ACTGCTTCGCGACCGGCACCGAACAGCCGGACGACAACTTCCTAGGCGAGCTGTACTACAGGG  
AATCGCGCATCGCCGGAAGTGGGACCGATGAGCCGCAAGCGGTTCCGGCTGGCGCGGGGTGGA  
GCCTGGCCGGCGACACCGAGCGTGCCGCGAGCTTCGCCGCACGCGCCCTGGCCGAAGGGTGGG  
CCGGCCACTCGACTCTGGCCACCGATCCGGATTTCCGCCGATTGCGCGCCAGCGCCGCCTGGC  
CCCGCCTCAATGTCCCTTGAAAGGTCACGTGGACTCATGATGTCCCCTTGAAAGGTCACACTC  
CGAGTCATGATGATTTGTCACTCCCACCGCTTCATTTTCCTCCACGTTCCCAAGGTCGCCGGC  
ACAAGCGTCAAGGACGTCCTCGGCCAAGAGCTATTCCAGGAGGACCAGGTCACGTTCCAGATC  
GCTCCCAATCCCCACTACCCACCTGAATGGACTGCGCCTTACGAGGAGCACATTATTGCCGCT  
GAATTGAAGAGCCAGTTGGCGCCGAAATTTGGGACGATTACTTCAAGTTCGCCTTCGTGCGC  
CATCCGCTCGACTGGGCGGTCTCCAATTACTTCTTCTTCCCTGCGCGACCGCAAAGGCCATCCG  
GCCACGAATTCCTGGAGCGGAAGGGCTTCGCCGGTACCATGGACATGTTTTTCGGAGCGGCC  
GGGCGCCATCCGCTGGTCGCCGGCATGCGCTTCAGCCAATGGGAGTTCTTGTGCGACAGCGAG  
GGCCGGACGCTGGTGGACTTCGTTGGCAAGTACGAGCGGCTCGAGCAGGACTTCGCCGCCGTG  
TGTATCCGCATCGGGCTGACCCCGCCCGACTTGCCGTGCCTCAACCAGACTCGCCACCAATCC  
TTTACCAGTTACTACGACGAGGCTTTGATGCGCCAAGTCAGCCGCGCGTTAGCTCGCGATTTTC  
GAAATTTTGTATTATGCCTGAGGCGGACCCGTTGCTTCGCCACCGGTGGATTATTCGATAAGT  
TATTATATTTTCAGTTGATCATGTGAATGTGATCCAGCCAACGAGGAGGATACCTCCGCGTG  
CGGCTATGGGGGCGCAGAGGTCACGACTACGTGTAGAAATTTGTGGAACACACCACTAGCTGC  
CACCGATTGGGAGCTTTGACTTGAAGATGAAAGTGGACAAGCGGAATGTGACGACATTTCTCG  
GACTCACTCCGACACAGACAGGCATCTTGTACCACTACCTGCTGGACCCGAGGCCGACGCCT  
ATTTGGAACAATTGACGCTGCACCTGGAGGGGCGCTCGACGTAGCGCGCTTCGCCGCGCCT  
GGGAGCGCGTGGTGGCGGCTCACGACCAGCTGCGCGCCGTGTTTCGCTGGCAAGGGATCGAAC

5

69

10

5

15

20

10

25

15

30

35

40

45

30

50

55

ACCCGGTGCAGATCATCCTCAAGCAGCACGTGCCGGACCTGGAGTTGGCGGAGGTCCCGCGCG  
ACGCCGATCCGGCAGCCTTCCTGGCGCAATGGGTGCGGGCCGACCGGGCGCGCAAGTTCGACT  
TCGAGACGGTGCCCTTTCGCATCGGCCTCTGCCGGACTGATACCCAACATCACGTGATGCTGC  
TCAGCAATCACCATATCCTGATGGACGGTTGGAGTACGGGCCTGATTCTGCGGGACTTCCTCG  
CCTGCTACGGCGACTCCGAAAACCTGGCGGCCACGCACCCGAACGCACTTCAAGGCGTTCATCA  
AGTGGCACCAGAACCGGCCACGCCGGGCGAGGAGCGATTTTGGCGCGACCTGTTGCGCGATG  
CGCCCGACGGCGGCTTTCGCCGCTGGGCGTCGAAGAAGGCACCCGCCACTCGCTTGACTTCG  
GCGCCCGCAGCCGCGCTCTCGACGACCGCTTGACCCAAGGCTTGCGCGACATGGCTCGCGACC  
TCGACGTCACCTCGCCGCGATGCTCCATACCGCTTGGGGCCTTCTACTCCAGCGCTACCAGA  
ACAGCTGCGAAGTGATATTCGGGACCACCGTTTCGGGCCGCAACGTGAGCTCGCCGGCCTCG  
ACGAGGTGGTCCGCTTGTTTCATCAACACGATTCCGTTCCGCTTCTCGGCCGCGGCCGCGACGA  
CGCCCGTCGAGGCCTTCCGTGCGGTACAGCGCAATCTGCTGGCGAGAAGCGAGTTCGAAGCCA  
CCCCGCTGGTGGACATCAAGGGCTGGAGTGGTCTCGGTCCGGGCGCGGAACCTGTTGACACCA  
TCCTGGTCATCGAGAACTATCCCTTGGACCGCGCTATCTTCGAGAGTGATTCCAGCCTGCGGT  
TGACCGACCACCAAATCTTCGAGCGCACCAATTACGGGCTGACCCTGACCATCGAGACCTTCA  
GCCGTTGACGCTAGCCCATCGCCGTGACCTGCTGGGCGACGCGGCCGCTGAGCGAA  
TGCTAGATCATTTCACCGGCCTGCTCCAAGCCATGCTGCGCTTCCCTCACCAGCCGTTGCGCGC  
GCCTCGAGATGAAAAGCGAACACGAGGCCACCGCGTCCTGCACCAACTCAACCAAACGCGTC  
AGCCGCTGCCGTCCCAATCGGCTTTCACACAGTTGTTCTTCGAGCAGGCCAGGCCGATGGGG  
CACGACCGGCGCTGTGGTGCGGCGCCACGCGCTGGACCTACGGCCAGCTGCTGGAACGTGCCC  
TGCGTCTGGCGGGACGGCTGCAGGAAGCCGGCTTCGCCCCGAGGCGATGTCGCCGCCGTCAGCC  
TCGGCCCCGTTCCGGATCTGATTCCCGGTTTGCTGGGCCCCGCTGTTGCGCGGCGGCGCCTACC  
TGCCGCTCGATCCCACCCTGCCGGCCAGCGCTCGCGGTTTCATCCTCGACGATGCCGTTGCC  
GCTTCCTGATCAGCGACGCGCCACTCGCGGGGCCACGCCGATCCATCCGGACCCTGCCGGCG  
CCAGCCCCGTTGACGTCATTTTTCCTGTCAGGACGGCGCCGCGCAGCCCCGCTACCTGATCT  
ACACCTCGGGCTCCACCGGCCAGCCCAAAGGCGTCTGGGTTAGCCACCGCAACCTGATCAACT  
TCCTGACGGGCATGAGCGCAATCCTGCCGGTTCGCGGCCGACGACGTGTTCTCTCGCTGACTA  
CCGTGTCGTTGACATTTTCGGGCTCGAGACGTGGTTCCCGCTCAGCCGCGGCTGCACGATCG  
TCTTGGGCACGCGCGCCGAGCAGTTGGACCCGGCCGCGGCTGCCAAGGCCATCTCCTGCCATG  
GCGTCACGGTTTACCAGGCGACGCCATCGCGACTCCAACCTCAACTGGAGCACCCACATTTG

5

70

10

15

20

25

30

35

40

45

50

55

TCCGCGCCATCGGCTCCCTGACGACCCTGCTGGTAGGCGGCGAACCCCTCCCAGCCGAGCTGC  
TGCGGCGCGTACGCGAAGTGACCGATGCGCGTATCTTCAACCTCTACGGTCCCACCGAAACCA  
CCATCTGGTCCACAGCCGGGAGGTACCGCGGCGGACGTCCCGGATATCGGCCGCCCCGATCG  
CAAATACCGGCGTTTTCTTCTGGCGCGAGACGGCTCGATCCAGCCGCGGGCCTGGTGGGCG  
5 AGTTGTGCATCGCCGGCGAGGGCGTGGCGTTGGGCTACCACCGACGGCCGGACCTGAACCGAG  
AACGGTTTCGCGAGATTCCGCCGGGCGCCTGCCCTTTGCCGGAAGCTCTACCACACCGGCG  
ACCTGGCCCCGCTGGACCGAAGACGGACGGCTCCTCTGCCTGGGCGCTCTGGACGACCAGCTCA  
AAGTGCGCGGCCATCGCGTCGAGCCGGGCGAGATCGAGGCAGTGATGGCGCGCCACCCGGCGG  
TCACGCAGGCGGTGGTCGTACGCGGCCGCGCAACGGCGAGCCGGTCTTGGTGGGTTCTGGA  
10 CTGCGGAAGGTGAGCCGATGCCAGAGGAAGCGCTGAGCGCTTACCTGGCCGACCGACTGCCGA  
GCTACATGGTACCCGAACGGTGCATCCTCATGAAGGCCATGCCGCTAACCGGCAACGGCAAGA  
TCGACCGGCGCGCCCTACCCAATCCCTTCGCCTTGACCGAGTCGACCCGGCAGGCGGCGCCGC  
GCACCTTGCCCCGCACCGCCGGCGAGCATCGGGTTGCCGAGCTGTGGCAGGCCTTGTTGCGAC  
25 GCGAGGCGATCGGCTTGGACGAACCCCTTTTTTCAGGCCGGCGGGAACCTCATTGGGCTTGATT  
GGCTTCACGCCAAGCTGGAATCCGCCTTCGGGAAGTCGTTCCCGATCACCGATTTGTTCCAGC  
ATACCAGTATTTCGAGCCAGGCAGAAATGCTGAGCGGCTCGTCCGTCGAGGCGCCGCTCGCGG  
30 GAGCCGTGCCGCAACCCCGGCCCGCCGCGCCCAAGTTGCCTCCTCGGCAGCTAAATCCCCAG  
GGGAGCGCGGCGCGGCAGCGACGTGAGCGGCCTGACCGCGCAACCGCCCCAACCCCACTTCC  
GGCCCATCGCCGTTATCGGCCTCGCCGGCCGATTCCCCGCGCACCCGACCTCGACGCCTTCC  
20 TTGAACTGCTCACGGAGGGTCGCTGCGGCATTGCTTCTTCAGCCAAGCCGAGCTGCGCGACG  
AGGGTCTCGACGCGAATCGAATCGCGTGTACATACTATGTCCCGGCCAAAGGTTTCTCGACC  
GGGCCGACCACTTTGATGCCGACTTCTTCGGCATCCCGCCGCGCGACGCAGAAATCACCGATC  
CGCAAATTGCGCTTCTGCTTGAGTGCTGCTGGAACGCGCTGGAGCATGCCGGCTACCCGCCCCG  
40 GCGGCGGCGAGATCGGGCTCTTCGCCGGCTCCTCGGCCAACTATCACTGGCTCGAATACGTGG  
25 GCATTTCCGAGGAGAGCAGCAATCGATTGCGCGTCATGATTCAAAACGAAAAGGACTACCTGG  
CCACGCGGATCGCCTACCAGCTCGATTTGAAGGGCATTGCCGTCACCGTGCAAACGGCCTGCT  
CGTCGTGCTGACCGCGGTCGAGCTGGCCTGCGATGCGTTACACGCCGGCCGCGTGACCATGG  
45 CTTTGGCTGGTGGCGTTGGTCTGACCTATCCGTTGCGCGCCGGATACCTGCACGAGGATGGAA  
TGATCTTCTCCCCGACGGTCGGTGCCGGGCTTCGACGCCAGGCGGCCGGCACGGTCTGCG  
30 GCAACGGTCTGGGCATGGTGGTGCTGAAACAGCTCGACGCGGCGCTGGCCGACGGCGATGCCA



5

71

10

15

20

25

30

35

40

45

50

55

TCCACGCTGTGATTAAGGGCATCGCGGCCAACAAACGACGGCGCGGCCAAGATCGGCTACACGG  
CGCCCTCGCAGAACGGTCAGGCGCGGGTGATCCGCGCCGCCCATAGGCTCGCCCAAGTCGCGC  
CGGAGACCATCGGCTATGTAGAAGCCACGGTTCGGGCACGCCGCTGGGCGATCCGATCGAGG  
TGGCGGGCCTGACCGAGGCCTTTGACAGCCCGCGTCGCGGCTTCTGCGCCTTGGGTTCGGTCA  
5 AGTCGAATGTGGGTCAATTTGGATGCGGCAGCGGGCATCGCGGGTTTCATCAAGGCGGTGCTCT  
CGCTGTCCCATCGGACCCTGTTCCGCCAGCCTCCACGTCGACACGCCCAACCCGCAGATCCCGT  
TCGCGGACGGTCCGTTCCAGGTCAACACGGAGACCCGGCCCTGGCCAGCTGCCGACCATCCCC  
GCCGCGCCGGCGTCAGCTCCTTCGGCATCGGCGGCACCAACGTGCACGCCGTCCTGGAAGAGG  
CGCCGCAGTTGGCCGAGCACGCGGGGCGGCGGCGGAGCGGCAGCTGTTCCCTGGTCTCGGCGC  
10 GGACTGCAGCCGATCTGGAGCGACGCACCGCGGCGCTGGTCCGCCACCTGGCCGCGCATCCGG  
ACCTCGCACCAGATGACGTTGCCTTTACCTTGCACGCGGGCCGCAAACCGATGACCCACCGTC  
GTTTCCTGGTCGCGCCCGACCTCGCGGAAGCCGCGCGGTCTGGCCGAGCCCGATCCAGTCA  
AATCCGCGCGGCGCGCGCCGACCGCTGCCAGGTCTGGATGTTCCGCCGTCTCGGCTCTCAAT  
ACCCCGGCATGTGTGGCGGCCTCTATCGCACCGAGCCGGCCTTTCCGCGAGCAAGTCGACCGCT  
15 GTTTCGACCTCCTCGCGCCGCGTTGCGATTTGAAGCCCTCGCTCTTCCCCGAGCCCGATCAGG  
CCATCGACGCATCAGCCCTCGCGGCCATCGACACCGCCAGATCGCCGTCTTCGTCTGCGAAT  
ACGCGCTCGCACGGATGCTGGAAGGCTGGGGGCTGCGTCCGGATCGGCTGATCGGTTACAGTT  
TCGGCGAATACGTGGCCGCCTGCCTGGCCGCGCTCTTCTCCCTGCCCGACGCCTTGGCAATCG  
TCCGCGAGCGTGGCCGGATCCTGGCGGCGGCCGAGCCGGGCGCGATGGTCAGCGTGCCCTTC  
20 CGGCCGAGCGCGTCGCGTCGCTGCTGGAGCCGCCGCTTGCTTGGCCATTGACAACGGCCCCCT  
CATGCGTGGTGTCCGGGCCGGTCAACCGGTGCGCACCTTCACCGCTCGCATGAAGCGGGACC  
GGGTCTGGGTGACGCCGCTCCAGGCCGAGCGCCCGATGCATTCCGCCGCTGATGGCCGAGGCCG  
GCGGCTCACTGCGCGCCATGTTGGCCGGGTTCCGCCTGAATGCGCCGCGAATCCCGATCTTAA  
GCAATGTTACAGGAACCTACCTAACCGACGAGCAGGCCCGAGACCCCGATTACTGGGCCCCGTC  
25 ACCTGTGCGGCAACGTTTCGCTTCGCCGACGGTGTGCGAACCTTGTTGGCCGAGCGCGATCCGG  
TGTTCCCTTGAATTCGGGCCGGGCCGCGATCTGAGCTCCTTGGTGCGCCACCAGATGCCGGAAG  
GCGCCGACGAGCCGATCGCACTGATCCGTCACTCGCGAAGATCCGGTGCGCGACGAAGACCTCC  
TGCTCGATGGCTTGGGCCGCTGCTTCCTGCGTGGGGCGACCCTCCACGGGCAGGCCTTGTACG  
CCGGCCGAGGCTGCCGCCGCGTGCCGCTGCCCGGTTACCCGTTCCAGGGTCCACGCTGCATGC  
30 CGGCCCCGCGCCGACTGCCCGGCCTGGCGCGACCGACCGTGGGAGCGACCACCATCAGCTACC



5

72

10

15

20

25

30

35

40

45

50

55

GACCAGCCTGGAAGCGGGCGCCGCGCTTGGCGGCTGTCGAATCGCTCGCGCCGCAATCCTGGT  
TGGTATTCAGCGACGGCAGCGAATTGGCGGGCGAGCTGGTGGCCGGCCTGCGCGCTTCCGGTT  
GCGCGACCACCCTCGTCGAAGGTGGGCTGGCGTTCGCGCGCTTCGCGGGCGGCTTCCGCGCGA  
ATCCCCGCGAGGAACAAGATCTCGCACAGCTGTTGCGGACCCTGTCGGCCGAAGCGATGCTGC  
5 CCACCCACATCCTGCACCTGCTCAGCCTGCCGTGCGCGGAGCGCGACTCGCCGCTGGCGCGCC  
TGGAGCACCTCACCGAGCTGGGCTTCCACCATCTGCTGGCCCTGGCCCGCCAACCTGGAGGCGG  
TCGGCGCCCCCGAGGTCCGCCTCGCCGTGGTGACAACCGGCCTGGCGGCGATTGGCGGCGAGT  
CCGAGCTGCGGCCCCGAGGTGCGGCTGTTGCGGGGACCTGTCCGCGTGATTCCCTTTGAATTCC  
CGAACTTGCGGCTGCGCCTGATCGACCTCGACTCGGCCGATCCCATCTGGCGTAGCGGTTGTG  
10 AGCCGTTGCTGCGCGAAATGGGCGCTGCCCCGGGACCTGAAGAAATCGCGCTGCGCGGCACCA  
GCCGTTGGGAGTTGGGCTACGAGCCGGTCGAGGGGGGCACCGTGAGCACCATCTCCTCGCGAC  
TGCGCGAGGGCGGCGTCTATCTGATCACCGGTGGCCTCGGCGGCCTGGGTCTGGCCTTGGCCC  
GTCACCTCGCCCCGGAAGTACCGCGCCACCCTGATCCTCGCTGGCCGGCGAGGCGCGCCGGCGC  
25 GCGAGCTCTGGCACCAGGCGCCAGCGGAGTTCGTACCGGTGCGAGCTGCGATCGCACAGATGG  
AGGAGTGTGGCGCCCGCGTGATTCCCGTCGCGCTCGACGTACCGACGCCGACCAAGTGAACG  
CGTTGTTGCCACCATAGAAGCTACGGTCGGCAAGATTGAAGGCGTTTTCCACATGGCTGGCA  
TCGTTGACGGCGGCATCATTGAAACGCGCACGCGCGCTGCCAGCGACCGCGTGCTGGCGCCCCA  
AAACGGTCGGAACCTGGATTCTCGATCGGGCTCTCCGCGGCGCCGGTGGCCGCTTCTGGTGC  
TGTACTCCTCGATCAACGCGGTGCTCGCGCCCTTCCGGCCAGGTTGCCTACGCCGCCGCCAACG  
20 CCTTCCTCGACGCCTTCGCCAGCGCCACGAACACGACGAGCGTCTTTTCCGCGTCAGCATCG  
GTTGGGACACCTGGCGCGAGGCCGGCATGGCCGTGATGCCGCCCGCGCCCGCGGCGACCAGG  
CCCCGCTCGAAGGGCTTAGCGACGAGCAGGGCTTGCGCCTGCTCGAAAGCGCCTTGGTCGGTT  
GCGAACCGCGACTCCTCGTCTCCATCAGCGAACTGCGCGCTCGACTAGCCGAGCATCATCGCA  
ACGGCGGCATTCCCCGGTTGCTCGGGCCCCCGCGCCAACGAGGCGGGTGCAGCTGATTCCGGCG  
40 AGGAGGGCGCCACGCAAGACGCGTCGCCGGCCCGTCGCGCCCGTCCCGATCTGGTCGTGGCCT  
TCGCGCCCGGCCGCAACGAGCTGGAGCGCCGGATCGTGGCCATCATCGGCGCCTACCTGCGGC  
TCGGTCAGGTGGGCGTCGACGACAACTTCAACGATTTGGGCGCCACCTCGCTCGACCTCATCC  
45 AGATCGCCCAACGCCTCGGTGCGGAGTTGGGCGCGATGTCCCTGTGCTCTCGCTCTACCAAC  
ACCGCACCGTACGCGGGCTGAGCCGCTTCTCGGCGGCGCGCTCCAATCCGCGCGGTCCGGCG  
30 TCCCGACGGGCGCTGCCGCACCGGCGCGGCCACGCGGGGGTTGCCACCCCGCGCGGCCAC

5

73

10

15

20

25

30

35

40

45

50

55

AACCGTCGCGCCAGCACCTGGAAAAACGCCGTCAATTGAGGAAAAAGGGGGGCCTTCCCATC  
ATGAGTGAAGTATCCATTGCCCCGGCTTGGACATCGCGGTCATCGGCATGGCCTGCCGCTTT  
CCCGGTGCCCGCAACCTCGCCGAGTATTGGGCCAACCTGATCGAAGGCCTCGAAACGCTCAGC  
TTCTTCAGCGAAGAGGAGCTGCGTGAGGGCCGGCTGCGATCCGGTCCAACCTGGCCCAGCACAAC  
5 TACGTGCGCACCAAGGGCCTGCTCCCTGACGCAGACCGTTTCGACGCCGATTTTTTTTGGTTAT  
TCCCCGCGCGAAGCCCAGGTGATGGACCCCCAGATCCGCGTCTTCCACGAGGTCTGTTGGCAG  
GCGCTGGAGCACGCGGGCTACAACCCGCATCGCCACACCGGCACGATCGGCCTGTTGCGCGGC  
GCCGCGCCCAACGTTTTTTGGGAGTTTCTCTCCTATCGGTCCGATGCCGCCAATTTAGGCAAC  
TTCACGCTGGGCCTGCACAACAACAAGGACTACCTGAGCTCGCGCATCGCCTACAACCTTCAAC  
10 CTGACAGGGCCCCAGCTACACCCTGTTACCCGCTGCTCGACCTCGATGGTCGCCATCCACCAG  
GCCGTCCAGGCGCTGCTCAACGGCGAATGCGACCTGTGCATGGCCGGCTCGGTCTCCATTACG  
CTGCCACTGGTTGCCGGCTACACCTACACGCGGGCATGATCGTCTCGCCCGACGGCCATTGC  
CGCACCTTCGACGCAGGCGCCAATGGCACTGTCTACGGCGACGGGGCCGGCGTGGTCGTTCTC  
AAGCGGGCCGAGGATGCGTTGGCCGACGGCGACCACATATTTGCGCTCATCAAGGGCTCGGCG  
15 CTCAACAACGATGGCAGTCGCAAGACCGGCTACACCGCGCCCAGCGTGCAGGGGCAGGTGGAG  
GTGATCCGCGCGGCGATGAACCTGCGCGAGGTGAGCCGGAGGCGATCAGCTACGTGGAAACC  
CACGGGACGGGCACCACGGTGGGCGATCCGCTGGAGTTCGAGGCGCTAAAGGAGGCCTTCGGA  
GGTGGCTGCAAGGCCTTCTGTGGATTGGGTTCGGTCAAGCCGAACATCGGCCATCTGGACGTG  
ACGTGCGGGATCGCGAGCTTCATCAAGCTGGTCCTGGCGCTGGAGCACCGCATCCTACCGCCC  
20 ACGCTCCAATTCCAACCTGCCCAACCCGAAGATGGATGTGGTCGATAGCCCCTTCTACATCGTG  
GCTGAGCGCGAACCCTGGCGCGAAGATCTGCTGCCGCGTCGGGCCGGTGTGAGCGCGTTCCGGT  
CTGGGTGGCACCAACGTCCACATGATTTTGGAGGAGTTTCAGCGCGAACC GGCGGCGAACAGC  
GCGCGCACGCGCCACCTGACGGTGCTGACGGCGCGGTGCGCGCAAGCCCTGGCGCAGCTGGCG  
GCCAACCTCGCCGAACACCTGCGCGAACACCCCGAGTTGGCGCTGGCCGATGTGGCCCATACG  
25 CTGCTGCACGGCCGCAAGCCACATCCATTGCGCGCATCCTGGTGGCGACCGATACGACGGCG  
GCGATCGACGCCCTTGATGAACGACCGCGATCCGCGAACGCGTTTCTTCGAAGCGACCGGGCGC  
GGCGAGTCGGTGATCCTGTGTTTTGACGAAACGCCGCCGAGCCGCGAAGCGCCCGCTACCTC  
45 TGGGATCACGAGCCGCTTTATCGC3CGGCGGCGACGTGCTGCTGGCTGGTGAGGTGCGCCGAC  
CCGGATCTGGAAGGCTGCTTTACT3CCCTGATCGCCGAGCAGGGCGCGGCAGCCGCCTTTTGC  
30 CACCAATACGCGCTGGCCGGATG3CTGCTGGCCATGGGGTTGACCCCGTCGGCGTTGATCGGC

5

10

15

20

25

30

35

40

45

50

55

GTGGGCCAGGGCGAGTGGGTAGCAGCGGCGCTCGCGGAGGTGTTCCCGCCATCGGCCTGCTTG  
CGCTGGATTAGGTTTCGGCGAACGGCTCCCGCAGCCGCGCGATCAACGGATTCCGTTTCTCTCC  
AATTTCTCTGGAAACTGGATCGTTGGGCGTGAGTTGGCCGACCCGGATTACCCCAGAAAGCAG  
AAGGGTAAGCGCTGCATGAAGCGCCGTCGGTCCCAACCTCGGTCAGCTGGTGCAGGATGGGGG  
5 CGATGGAACCGGCTCGGTCAGCTCGTCGCGCGCTGCTCTTCCGCGGGAAGCGGAGGCGGGACG  
GTGATCGGCCCCGAGGGCGAGGTTTCATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCCAG  
15 TACCTGGGGGCGAGCTCGAGGTAGCGGTCCCGCGGCCAGTAGGGCATCGCGCGAATGACGTCG  
GCCAGGTAGGCCTCCGGGTGAGCCCCGTGCAGCTTGCAGCTCGCCACGAGCGAGAAGAGGTTG  
GCCGCGGCGGAGGCGTGCTCGTCGCTGCCGAAGAAGAGCCAGGACTTTCTCGCAACCGCAATG  
20 GATCGCAGCGCTCGCTCGCTGGCGTTGTTCTCCAGGCGCAGCCGACCGTCGTCGAGGAAGCGC  
CGCAACGGCTGCTCTTGTTGAGGGCGTAGCCGAGCGCGGTGGAGACCAGGCCGCGCTCGCGG  
GGACGAGCGTGCTCGGCCCTGGCCCAGGCAAAGAACGCGTCGACCAGAGGGCGGACGACGACA  
25 TCGCGACGCACCTTGCGCTGCGCGGGCGGCAGGTCCGCCAGCGCGCGATCGGCGGCAAAGAGG  
GCGTTGATGCGCCGAGCCCCCTCGACACCGAGCTCGTGCTTGCAGACCGCCGCTCCAGAAAG  
15 TTGGTACGGCAATGCGACCAGCATCCGACTTCGGTCGGGGGCGGACCGCGCTTCTCGTCGGCA  
GCAGCGCCTCTTGTTGGTGTGCCGCGGAAGAGGGCGTCATAGATGGCGTGAGCGTCAGCTTGA  
30 ATATACCGAGAGAAGCCGCGGAACATCTCGCAGACCGCGGCGCTGGTATGCTTGGGCTGGTAC  
TCGAAGAAGACGTGATCCTTGTCGCGGAGGACGACGAAGAAGTGTCCCTTGCGGCACGGCCCCG  
GGCTTCTTGTCCTTGCGCTCCTGGATGGGCCCAGGCTGGACGGAGACCCCGGTGGCGTCCGTG  
20 GACAGGCAGAAGGCGGTCTCGAAGGCCTCTTTGCGCGCGGCCCTCGACGATGGCGCCCAGGGTC  
35 GCACCGACGTCTTCGGCGTAGCGGCACATCGTGCCGCGATCGAGCGACGCGCCCTGAAGCTCC  
AGCTGCTGCTCCAGTCGATAGAACGGGACGCCGAGCAGGTACTTGCTGGTGAGGATGTGCGCA  
ATCATCGACGGCGCGAGGAACGACCGCCGGAACAACCTCCTTCGGAAGCGGCGTCGTGATGAAG  
40 ACCGTGCAGGTCTCGCCCTTCGGCGCCGGCGGGCGCGCTCGAGCGAGGGCGCGTCGAGCGCT  
25 GTGGAGGAAGCGCTCGGCTCGCCGGCCGCTGCCGTGTCCTCCGGGCTGACGCTCGCAGCGGGC  
GTCCGCGTCGGGGCTTCTCTCGCGACGACCTGGAGCGGGGCCGCTTCCTCCTCGCCCGAACTG  
CTCGCATCCGTGACGGACCGCTCGGCCTTGTAACGACGCGTGCGAGCACGATGCGGCGCATT  
45 CCGCCGCGCTCGTAGCCGAGTCGCGAGGTCTCCTCGACCCCGATGCGCGTCGCCGTGCGCATCG  
AGCTCGGGGCGAGGAGCTCGATGCGGACGACGGGCGAGTCCGACTCGGACAGGTGCGGACGG  
30 CCCTTGCCGCCGGACCTTCGTTTCGGCCCCCTTGGGGTCGTGCTGCTGCCGCTCGTCGCCTGTA

5

75

10

15

20

25

30

35

40

45

50

55

TTGCGCTCGGCGGCGTCCGAGTGCCCTTCGCGAGGCGCTGGACCTCGAGGAACATCGAGTCGAAC  
GCCAGCTGCTCCGCGCTCACCTCGGCGCGCTCCGCCTTGGCCACGAACAGTCGACGTCGCAGA  
AGCTGCAGCTGCTCGAGCGCACGGGTGTAGGCGCGCCGAAGCTGCGCGAGCGCATCGCGCGCT  
CCCACGAGCTCGCTCTTTGCCGCGGCGAGCTCCGCTTCGAGCTGCGCGATGCGCTGCTGCTCG  
5 GCGGAGAGCGTCGGCTTGGCGGCGGCGTCGTGCACGACGCCGCTCTACGTAAGCCGCGCGTAC  
TTGTCGAGCGAATTTCGTGCGGCTCAGTGGACGCGGCGCGGTGCGCGCCTTCGCGGTTTGGACG  
15 TGGGCGCGATCTCGATGCCGTGAGCAGCGTCTCGAGCGTGGCGTCGTCCACCTCGACGTGCG  
TGGCGCCCTCGGTGCGGGGGTGGGAAGTGCGAACGCTCCGCGATCAAGGCGTTTTGAAAACA  
GGCAGATTCCACTGCCATCGAAGAAGAGAATCTTGATCGTGGTCCGCCGCTTGCCGACGAACG  
20 CGAACAGCGCTCCGCAGCGAGCCTCGTACCCACACGCTCACGGATGAGACCCGAAAGCCGCT  
CGAAGCCGTAGCCCATGTCCACCGGCTCCAGCGCGACGAACACCTGCACGCCCCGCCGAATCA  
TCGCCCCGCTCCGCCGAGGGCACGGACCACCTCCGCCAGCAGCGCGGGTTCGAACCCCGCGGC  
25 GACGCGACCCGCGCGCCCGACCTCGACGACGAGCTCCGCAGCGCTGCTCGTCACGGCGGG  
CGCCTTCGGCACCAGGCGCAGAAAGCGCGGTGGCTCGGCCCGCGACAGCCGGCTCGACCAGCC  
15 GTGCAGCGTCGAGGCCGCAAATCCGCGGCTCCGAGCGAACTCCTCCGCCGTTTCACCACTCTC  
GCGCCACGCCCCGAACGCGCTCGGACCACATCACTTCGGTCGCCTTCGTCTTGTTCATGCACGC  
30 CATCATGAACTGGACAGCGCAGCCGGGTGAGACGGCGCTTCGCGCAGCGCTTACGCAGAAGG  
CGCGCCGCGCGCCATTGTTCGGATGCGGTGCGCGACTTCGCCGCCGATCGGCTGTTGCTGGAAC  
TGGGACAACCACTGGACGTAACGGCTGAAGCGAGCCAACGGCTCCAGCTCGCGCGGGGCGACC  
20 TGTTTCGGCGCCTACCAAGCGTTGGCCCAGCTCTGGATCTGCGGCGCCCTGGCCGAACCGCCGC  
35 GACTGTATCCCGACGAACACCGCCGGCGCGTGCCGCTGCCGAGCTACCCCTTCGAGGGAAAGC  
GGTTCTGGATCGAGGGCTCGCCGTTTCAAACCGCGCCCCGCCGCGCGCTCACCCCAACCCG  
CCGATTTCGGGGGACATTCTCAAGGGCGACCCGGCGGACTGGTACTATCGGCCGCGTTTCGAAG  
40 CGGCGCCGCTCTTGCCCAGCCCGTTTCGAGAGCGAACCCGGCGATTGGCTGGTGTTCGAAGATG  
25 AGCTGGGGCTCGGCGCCTGGCTGAGCGAGACCTTGCGCGACAAGGGCGCGGGTTCGCGACAG  
TCGTTTCGAGGCACCGAGTTCCGACGCTTGCGGTACAGCGCTTCAGCTTCGTCCCGATCGAC  
GGGACGATTACCGGACCCCTGCTGCACGAGTTGAAGGCGCAGGGCATCGCGCCGGTCCACCTGT  
45 GCCACCTATGGAGCGTGACCGCCGCACCGGATGCCGAGCAGTTGCTCGACGTCAGCTTTCACA  
GCCTGGTCCATTTGGCGGCCGCTTTGGGTTCGGTTGGCTACTTCCACGCCATGAAGTTGAACG  
30 TGGTCGCCAACCGGCTATTCGACCCCGAGTCGCCCGAGCGCACCGAGCCCGCCAAGAGTCTGT

5

76

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

TGCTCGCGGTGACCAAAGTCCTGCCGCAAGAGGTGCCCAACGTTCTGAACCCGCGCCATCAGCG  
TGGACCTGGATCGCTCGTTCGACGCGGCGGCGCCCGCCTGGGCGCCAGTTTGTGGTTGAAT  
GCGGCGCGCCCGTTCGAGGAAACGGTGGTGACCTACCATGGCGCAGCCCGATGGCTGCGCCGCT  
TCGATCGCGTTGCGGTGAATGGTCTCGGCCCCGTTCACCCCGATCAACCTGCGCCGCTGCTGC  
GCGAGCGCGGCGTGTACCTGATCACCGGCGGCCTGGGCGGCGTGGCTGGCCAGTTGGCGCGCT  
ACCTGGCGCGGCGCTGCCGGGCGCGGTTGGTGCTCACCGCGCGCCGGCCCCCTGCCCGAGCGCG  
ACCAGTGGGATCGGGAGTCGGCCGTGCTGTATGGGACGACAAGACGCGCCAGCGCATCGAGC  
TGGTGCGCGAGCTGGAGCGGCTGGGGGCCGAAGTATTGGTGGTGGCTGCCGATGTCGCCGACG  
AAGCGGCCATGGCGCAGGCGATCGAGGCCTCACTGGCGCGATTTCGACGCTTTGGACGGCTTGA  
TCCACGGCGCCGGGATCGTGCGGCTCGCGTCCGGCCGACGCGCGATCGGGAGTATGACGCGGG  
CCATGTGCGAGGAGCAGCTCCGCCCCAAGATGTTGGGCCTCGACGTCGTCGACCGCCTCCTGC  
GCGATCGCCGGTTGGACTTCCGCATTGCCATCTCGTCGCTCGCCCCGATTCTCGGCGGCCTCG  
GCCACGTCGCCTACGCCGCCGCCAACCTCTACATGGACGCGTTTCGCGACGCGCGCCGCCGCCG  
GCAACGCGCCTTGGATCGCGCTGAACCTGGCCGAGTGGGAATACGAGGGCCCCGGCTACCTACG  
ACGAGCGGGTGGGCCGTTTCGCTCAAGCAGCTCGAGCTCACCAACGAGGAGGGTATCCGCGTCT  
TCCAGACGGTGTGGCCTTGGCCGCGCGCGGCCCGCTACAGCAGATCATTATTTCCACCGGCG  
ACCTCCAGGCCCCGCTCGACAAATGATTACATCAAATCCCTGCATCGCCGACCGGGGCCGG  
TCCAGCTCAGTCGCCGGACCGCGGCACCCCAGGGCGGTTTCGGCTCGGAGCGCGCCGCCTTCG  
AGGCCGCCTTCGCTGACGCCTGGTTCGACTTCTTCGGGGTTGAAGAGGTTCGACCCGAACAAAA  
ACTTCTTCGATCTGGGCGCCAGCTCGCTCGACTTCATCCACCTCGTCAGTCGCTTCAGCAAGG  
CCATCGAACAGCATGTACCGCTCGAGGCCCTGCTCGAACACTCCACCCTGCACGACCTCGCCG  
CCCACCTCGCGGGCGACGCGAACACCGACGCCAGCGACGAAGCGCGCATTTCGCCAACGGCTGC  
AAGGCGCCAAGTCCGGCGACATCGCCATCATCGGCATGGCCGGCCGCTTCCCGCTCGCGCCCCG  
ACCTGGACACCTATTGGCGCAACCTGGTTCGGAGGCATCGACGCGGTCAGCTTCTTCAGCGCCG  
AGGAGTTGCGTGCTGCTGGCGTCAACCGCGGCCGAGATCCACCACACCAACTACGTGCCGGCCA  
AGGGGCGCTGCGCCGACCAGGACTTGTTCGATGCGGCCTTCTTCGAATACACTGCCAGCGACG  
CCGAGCTGATGGACCCGCAAAATCGCGTGTACACGAGGTCGTGTGGCACGCGCTGGAAGACG  
CCTGTTTCGACTTCAACGGCGATCACGGCCAGGTCGGCCTGTTTCGCGGGCGCCTCGCCGAACC  
TGTGGTGGCAGTTCGTGGCCAGCTTTTCCGAGGCCGCCAAGACGCAGGGCATGTTACACACCA  
CCCTGCTCAACGACAAGGACTCGATCGCGACCCAGATTTATACAAGCTCGGTCTAAAGGGCC

5

77

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

CCGCGGTACCTTGTTACCGGCTGTTCCACCTCGCTGGTAGCCGTTGACGCCGCCTGCCGCT  
CGATCTGGTCCGGTCAATCGGACATGGCCGTGGCCGGCGCGGTCTCGCTGACTCTCCCCGATA  
AGGCCGGCTACATCTACGAAAAGGCGATGCTCTTCTCGGCCGACGGCCATTGCCGGGCTTTTCG  
ACGCCAACGCCACCGGCATGGTCTTGGCGACGGCGCCGGCGCGATCGTGCTCAAGCCGTTGG  
ACGCGGCCCTGCGCGACGGCGACCCGATCCATGCGGTGATCAAGGGCTGCGCCACCAACAACG  
ACGGCGACCGCAAAGCCGGCTACAGAGCGTCAGCGCCCAAGGCCAGGCCGAGGTGATCCGCT  
CGGCCAGATCCTGGCCGACGTGGCGCCCGAATCCATCAGCTACGTGGAAGCCACGGTACCG  
GCACCAAGTTGGGCGACTCGATCGAGATCAAGGCGTTGAAGCAAGCCTTCGCCAGCGACAAGA  
ACGGATTTTGCGGCATCGGGTCGGTCAAGACCAACCTCGGTACCTGATGGCGGCGGCGGGA  
TGGCCGGCCTGATCAAGACGGTCTCGCGATGAAGCACCGCCAATTGCCGCCATCGCTGCACT  
GCGACGAAGTGAACCCCGACCTGGGTTGGAGCGCAGTCCGTTCTACATCAACACCCGCCTGC  
GCGACTGGGTTGCACCGGGCGGGCCCTGCGGGCCGGCGTGAGTTCGTTCCGGATCGGCGGAA  
CCAACGCTCACGTCATCCTGGAGGAGCCGCCGACGCGCGAGAGCGGCACGCGCATGCGCCACT  
GGAAATTATTGATGCTGTGCGCGGCCAGCGAGGCGGCGCTCGACCGCCAGGCCGATAACCTGG  
CCGACTACCTGGAGCGCCATCCCGAGGCCACCTCAGCGACGTGGCCTATTCCCTCCAGACCG  
GCCGGCGCGTTCTGGCCTGGCGGCGCACGGTCCTATGCGAGTACCGCGAGGACGCGGTGACCA  
GTCTGCGCGAGCGACAGGCCAAGCCCGTCCAGACAAGTCGCGTCCGCTGGGACCACAAGGACG  
TGGTCTTCATGTTTCCCGGTGAGGCGGCCAGTACCTCAACATGGGCCGCGACTTATACGTCA  
TGGAGCCGGTCTTCCGCGAGGTGATGGACCGCTGCTTCGAGTTGCTGGCCCCCTTTGTGGTCCG  
AGCATCCGCGCCAGATCCTTTATCCGAGGGCGGGGTGTGACCCCTGCTCCACCGGACTGATT  
ACACCCAGCCGATCGTGTTCTGCTTCGAGTACGCCCTCGCCCATTTGCTGCTCTCCTGGGGAT  
TGAAGCCGGCCGCGACCATCGGCTACAGCTTCGGCGAGTACGTTTCTGCCTGCCTCGCCGGCG  
TCTTCTCCCTGGAAGATGCGATCCGCTGCTGGTGACCGAGCGCGGTCCGCTGATGGCGGCTTTGC  
CCGCGGGCGCCATGCTCAGCGTCCCCTTCCGAATGCGAGCTGCTGCGGCTGCTGGACGGCT  
TCCACGCCCAATCGGCGGCCCATCTGGCGCTGGCCGTCGACAATGGCGCCTCCTGCATTGTGG  
CCGGCGAGCAGGCCGCCATCTCGGCCTTCGAATCGATGCTTCGCAAGAAGCGTCTGTTGACCA  
TGCGGTCGCGGTGAGCCACGCCGCTCATTCGCAGGTGATGACCGGCGCGACCGACGCCCTGC  
GCAGCATCCTGCGGAAGATCCCCCTCTCCGCGCCGACAATTCCCTTCATTTCTGCGTCACCG  
GCACCTGGATCACTGCACAGCAGGCTACGGATCGCGAGTATTGGGTGAACCACATGTGCGGGA  
CGGTGCGGTTCGCGGCGGGTCTGACCGAGCTGGGTCAAACCGCGAGGCGGTGTTCTGGAAG



5

78

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

TAGGTCCGGGCGCGACTTGACGTTGCTGGCCACCGCATCCTGGCCGACAGCGCGGCCGTGT  
TCGAGCTGGTCAAGGCGCCCGACGGCGGCGACGACGATGGGTTCCTCCTGCTGGATCGATTGG  
CCAAGCTCTGGAGGCTGGGGATTTCGATTGACTGGGCCGGCTTCTACGCGGATGAGCGGCGGC  
GGAAACTCTCGCTGCCGGGATATCCGTTTCGAGCGGCGGCGCTTCTGGATCGAGGGCAACCCGC  
TGGAGATCGCCGCCGGCAGGCCCAATGTCCAGGGGCGCTGGTCAAGGCGTCCGACATCGGCG  
CTTGGTTCTACGTGCCGCAATGGCGGCGGTGCGTGCTCGCCGAGCCGGGTACAACGGCGGCGG  
GCGCCGCGCTCACGGCGGAGCAGGCACGCGTCGTGACCGAGCTACGGGCGGGATGCGCGTCCG  
CCGGCTTGGGCAGCGGGGCTGCGGACTGAATGGCGGTGCCCCGTCCGAGCGTCCGAAAGAAA  
GTGTAGCGCCAGCCGGGTGACCCAGCGCAGCGGCGCAGACCGGCGCGGACTGCCCCGACACCGA  
CTGGGGAGCCAGCGGCTGTGCCAAAGGACGGGGCCGAGCCGCGGCCGACCTGGCTTATTTTCG  
CCGACGCCGGCGGATTGGCCGAATCTTTTCGCCAAGCGGGTTTCAGGCCCGCGGCGAGAAGCTTT  
ACCTGGTGGCTTCCGGCTCGCGCTTCGAGCGCCTGGCCGAGACCCGCTTCCGCCTCGATCCCG  
GGGCCAAGTCCGATCACCGCCTGCTTTTCAAGGCGCTCGACGAGGCCGACATCCTGCCGACCC  
ACCTCCTCGACTTCCGCTCGCTTGACTGCGGCGGGCCCGACGCCGACCCCATGGACCAGGCCG  
GCTTCTTCGGGCTGTTGCACCTGGTCCAGGCGATGGCAGAGGCCGGCTACAGCCATCCCATTC  
GGCTGCTGATCGTCAGTTGCGGCGTCTACGATGTCACCGGTGCCGAACCGCTGCAGCCGGCGC  
GGGCCACGATGATCGGACCGGCTCTGTGCATCCCGCAACAGTATCCGCACCTCGAAACGAGCC  
ATGTGGATTTGGGCGTGGTCCATGCCGACGAGCTCCACGCCGCGCGCCAGCTCGACAGCCTAC  
TTGCCGAATGCCTAAGTGCAACGGCCGAGCGCCAATTGGCGCTGCGCGGCCGACACCGCTGGC  
TGCTGGACTACGAGCCAGTCCGCTTGCCGCCGCTCGACCCGGGCCGTCTGCCCTGGCGCCAGC  
GCGGGGTCTACTTGATCACCGGCGGTTTGGGCGGGATCGGCCGCATCCTGGCCGAACACCTGG  
CCCGCACGACCTCGGCTCGCCTGGTCCTAATCGGCCGCGAAACCCTGCCCGACCGCGACGACT  
GGGACGCCTGGCTGAACCGCCCGCAACCGGTGACGCCACCCACGAACGGCTGCTGCACAAGA  
TCCGCGCGATTTCGCGATCTGGAAGCGCTAGGCGCCGAAGTCTTGGTCCTCGCCGCCGACGTCC  
CCAACGAAGCCGCCATGCGCGAGGCCTACGATCGCGCCGAATCCCACTTCGGCACAATCCACG  
GGGTGATTCACGGCGCCGGCCTGATGGACGCGCAAAGCTTCTCACTGATCGACGCCCTCGACC  
ACGACCTCTGCGCCCGCCAGTTCGAAGCAAAAATCCGCGGCGTCTGCGTGCTCGACCGCGTTC  
TGGCCGACCGCACGCTCGACTTCTGTCTGCTGATGTCTTCCATCTCCACCGTGCTCGGCGGCC  
TGGGCTATTTCCGTTACGCCGCGGCCAACGCCTTCCTCGACGCCTTCGCCAGGCGCGCAGCC  
GCCACGCCGCTTTCCCTGGCTTAGCGTGGCCTGGAGCGATTGGAAGTACTGGACCGAGCGCA

5

79

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

AGATGGACAACGAGGTCGGCGCCGTCATCGACAGCCTCTCGATGGAACCCGCCGAGGGCTTCG  
AAGCCGTCACCCGCGTCTTGGCTTGGGGCAAGGCGCCCCACATCGCCAACTCGCCCGGTGACC  
TCGGTCGCCCGCCGGGATCAATGGGTCAAACCTGGCCAGCCTGAAATCGGCGCACTCCAGCGAGC  
CCGAGCCGGCTAGGCATGGACGTCCGGCGCTCTCCAGCGAATGGGTGCGCGCCGCGCAACGTGG  
TCGAAGAGAAGCTGGTCGCCATTTTCGAGCAGGTGTTCCGGCACTGCGGCACTGGGCATCGAGG  
ACAACTTCTTTGAGTTGCGCGGCGACTCGCTCAAGGCGGTGATGACCGCGGCCCCGTATTCAA  
AGGAGCTGAACGTGGAAGTGCCGCTGCCGACCTTCTTCCAGATGCCACGGTCGCTGGCCTGG  
CCCAGTTCGTGACGCAAGCCAAGCGCAGCGGCCGGGAGACGATTCCGGCGCACCGCGCCGCGCC  
CACATTACCCGCTCTCGGCTGCCCAGGGCCGCCATTACCTGCACTACCGCATGGACCCGCGTT  
GTACCGCATACAACGATCCCTTCGCCAACCTGATCGAGGGTCCGCTGGACGTGGATCGCGTGG  
AGCGCATCCTGCACACCCTCATCCTACGCCACGACTGCTTCCGCACCTCGTTCCACTTCCGCG  
AGGGCGAGCCGGTCCAGGTGATTACGATCGGGTGGACTTCAACCTGGCGCGGATTACCTGCG  
CGCCCGAGGATTTGCCCCGAACGGATGCGCGATTTTCATCCGCTCCTTCGATCTGGAGCGACCGC  
CCGCCATGCGCGCCGGCCTCTTCGTACGGGGCCCGAGCGCCACGTGCTGCTAATCGATTTTC  
ACCACATTATCACCGATGGCGTGTGCTTCGAGAACTTCGTCCGGCGAGTTCGCGGCGCTCTACC  
GCGGCGAGATCCTGCCCCGAGCTGGAACCTCGAGTACAAGGATTTCCGCGGTGTGGCAGCATGAGA  
ACCGGGGCGCCGCGCAACAGCGACCAGGCCCGCTACTGGACCGAGCAGTTGGCCAATGCGC  
CCGGGCCGATCGAGCTAACCACCGATTTCCCCCGTCCAGTCGACGCAGCTTCCGCGGCGACC  
GCGTGCGGACCGTGCTTGATGCGGAGCTCGTTGCTCGACTCAAAGAGCACGCGGCGCGCCTCG  
GCATCACCTCTATAGCCTGCTGCTGGGCGGATTCTCGTTATTGCAGCACAAGCTCTCCGACT  
CGCACGACATCGTCATCGGTTCCGCCGTCGCGGGCCGCACCCGGAGCGAACTCCAGGATCTGC  
TGGGCGCGTTTCGTCAACACCCTGCCGATGCGCCACCGCATCGACCCGACCCATAACCGCACGGG  
TCTTCTTGAGCAGGTCCACCAGACAACCTTGCGGGCCCTCAGCTACCAGGAGCACCTTTTG  
ACGAAATGGTGGCGACGCTCGGGTTCCGCCCGGATCCGGCTCGCAACCCGATCTTCGACACGA  
TGTTCTTGCTGCAGAACATGGCCATGGGTGCAACCACCATTCGCCGTCTGCGGCTCTCGCCTC  
ACGACACTTTTCACCGCAAGGCATTGTGCGACCTGATGCTACAGGCGACCGAGTATGACTGCC  
ACCTGGAGCTGGTGCTCGAGTTCGCCACCGACCTGTTCCGGCTGGAAACCGCGCAAGTCTTGC  
TCGACCGCTACCGCCAAGTCTTGAGTGGCTGTTGGCGTACCCCATGAATCGATAGACGATT  
TGACGCTCGCCGGCCACTTTTCGCGAAGTCGAAGTGACGATGTCCGACGAGGGCGACTTTGATT  
TCTCAGATTTTGAACCCCGCAACGTGAGAAACCTATGGCGCGCCTGAGCCGCACAGATCTCCA



5

80

10

15

20

25

30

35

40

45

50

55

ACTCGCCATTACACGCGCACCGTGGAGCGCGAATATTGGCGCGCTCTGTTGAGCGCCATCC  
GCAACGGTCCAGTTTGGCGGGGTGCTCACCGCCCCGATCGGCGACGAGTCGACCCGCGAGAC  
CTTGTCATTTCGTCTTCGACGAAGATCCCCTTCGGCTGAGTAATCGTTTCGCCGCAACGCCTGCT  
CACGGTGTGGCGGCTGGCCTCGCGGCTTTCTCCACCGCTGCGACGGCGCTGAGCGCTTCAC  
5 CCTGGGGTTGGCCCTACCGCGCCAAGCCGATGACCATCACCCGATCCTCAACAGCTTGATCGC  
GCTGGGGGTCGCGGTGCACTCGAGTACGACCTTCCGCGATCTGCTCTATGCGCTTCGATCCGA  
ATACCACGAGGCGATGCGCCACGCCAACTTTCCGCTGGCGACCTGGTGGCGCGGCCTACCCGG  
CGGAACGGCGCCGTTTCGACGTGCGCCTCAGCCTGGACCCCTTCACAGACGGCGATTTCGCTGGA  
AGACCACGCGATCGGCGCGTTGTTCCGGTTCGCATTGGAGGGTGAGCGCCTCACCTGCCGATT  
10 GCGATTCGACCCTGCGCGCTATGACCGTCCCGCGATCGAAAACCTCGCCGATCGTTTCGCCCG  
CTTCCTCACGCGCCTGTGCCGGGACGCTCCACCGTCATCCAGGCGCTGGACCTTTCGCTGCC  
AAGCGATGAATCGGTGTGGCGCGTCACTGAAGGCGTGCGGCGCGGCTATTTCGCAAGACCTGAC  
GCTAGACCGCGCGTTCCGCCGCCAGGCCGCGCAAACGCCCGATCAGCCGGCGATCACGTTGAA  
CGGGGACGTCCAGAGCTACGCCGAGGTGACCGCCGACGCGCTGGCCCCGCCACCTCCG  
15 TCGCCACGGCGTCGGTCCGGAACGATTGTGGCCGTCAACGCCCGGCGCGGGCCTAATCAGCT  
GACGGCCCTGCTCGCGGTCCATAAGGCCGGCGGCGCCTACCTGCCGATCGATGCCGAGGAGCC  
GGCTGCCCGCCAGCAATTCAAGGTGCGCGACAGCGGGGCGCGGTTGGCACTGGAGCCGTGCC  
GGACCAGGCGCTGACCGTCACCGACCTGCCGCGGCTCTTCCTGGACGATGCCTCGCTCTTCGC  
TGACGGCGGGCTCGATGTGCCGCGCGGCGCCGACTCGCTCAATCCGGCCTATGTGATGTACAC  
20 GTCCGGCTCGACCGGACAGCCCAAGSGTGTGGTGGTTCCCCACCGCGGCGTGGTCAATCGTTT  
GAATTGGGGGCAGTCCCGTTTCCCCTGGACGAACCGGACCGAATCCTCCAAAAGACGCCGCT  
GCTGTTTCGACGTGTCCGTCTACGAGCTGTTCTGGGGCGCATGGAGCGGGGCCACCCTGGACAT  
CCTCGAGCCCCGGCGCCGAGCGCGACCCCGACGAGTGGCCAGGGCCCTGGCCGAGCGCGCCAT  
TACCGTATGCCATTTTCGTGCCTTCGATGCTGCTCGTCTACTTGGAAGTCATGCGGCGGCACCA  
25 TGCGCCGCCCCGTGCCGGACCGGCTCCGTTACGTCTTCGTCAAGTGGCGAGGCCCTCGAACC  
CCACCTCGCCGGGCTCCAGCAGATTGGTCGGCGCCTCGGCCGCACGATTCCCCTCGTTAATCT  
GTATGGACCAACCGAGGCCTCGATCGAAGTCTCCTGCTTCGCCTGTCCCGCCGACCATGTGCC  
45 GCGCCGGATCCCCATCGGGCAGCCGATCGACAACGTGCGACTGCACGTTCTCGACCGGCGCGG  
CCGTGCGCCAGCCGCCCTATCTTCCTGGCGAGCTGTTCTGGCCGGCGACTGCCTGGCGCGCGG  
30 CTACCTCAACCGTCCCGACCTGACC3CGCTCCACTTCGTGCCCAATCCCTTCGGCAACGGCGA

5

81

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GCGCATGTACCACAGCGGCGACTTGGCGCTCGTGCGCGGCGACGGCCAAGTGGCGTTTCTCGG  
CCGCCGTGACCACCAAATCAAAATCCGTGGTCAACGGGTGGAATGGGCGAAATCGAGAGTCA  
TTTGCGCGGGCTCGAAGGCATCGCCGCCCGCGTCGTCCAGGCCGAGTCGCAGCACCATGAAAC  
CCTGCTGCACGCCTACGTCTGTCACCAACGACGCGGGCCTCAATGCGGCCCCGGCTGCGCGCCGC  
CCTCGCTCAACATCTGCCCCGAGTACATGATTCCCCAGCGCTTCTCGCGGCTGGCCGAGTTGCC  
GCTGCTGGCGGCAGGCAAGATCGACCGCGCCGCCCTCGCGCAACGTGCAACGCCGCTCGCCAG  
CGCGCGCCCTTCTGTGGAACCCAGCGGGCCCCACCCAGCAGCGTATCGCAGAACTGTGGCGCCA  
GGTCTTAGCGGTGCGCGAAGTCGGCGCCGAGGATCCCTTCTTCAGCATCGGCGGCAACTCGCT  
CAATGTGCTCAAGCTCAGCGCCGCGCTGAGCGACGCTTTCGCGCGTGACATTCCCATGCCGCG  
CCTGTTCCAATACGACACCATCGCCGCCAGGCCTCCTGGCTCGACGGGCAGGTTGACGAACG  
GGCCCAATCCGCCGCGCTCGACCGCGAGGCCGCCGAGGCGGCGCTGACCCCTCAAGAGACCGT  
GGCCATTTTTGAGGGATTTCGATGACGAACCATGACCATCACGAGGAGAGCAGCGGCCTGGAGA  
TCGCCGTCATCAGCATGGCCTGCCGATTCCCGGGTGCTGCCGATTGCGACGCATTCTGGGAAA  
ACCTGATCAACGGGACCTCCTCGATCACCCATTTTCAGCGACGACGAGCTGATCGCGGCCGGCG  
TTGACGCGCGCGACCTGACGCCGCGTACGTGCGCGCGGCCGCCAGATCGATGACGCCGAAC  
GGTTCGACGCGGCCTTCTTTGGGTACTCCCAGCGTGAGGCCGAGCTGATGGACCCCCAGTTCC  
GCCTGCTCCATGAATGCGCCTGGTCTGTCTGGAACAGGCCGGCATCGATCCGCGCGTCAAG  
CCGCCCGATCGGGCTGTATGCCGCGCAGCCGACAACACCTACTGGAACGCGCTCTCGTCGC  
TCGACCGGGGCTCGGCCGAATCGGAGCAATTCGCCGCCGAACAACTTTGCAACCGCGATTTTC  
TGTGCACGCTGGTCGCCGCCGCGCTCAACCTGAAAGGCCCGCGGTGGTGGTTCAAAGCGCCT  
GTTGACCTCGCTGTTGGCGGTCCACTCGGCCTGTCTGCGCTCCTGACCGGCGAATGCCGAG  
TGGCCTTGGCCGGTGGGGTGGCGCTGCGCTTCCACGCCCCGAGCGGTATCGCTACGAACCTG  
GCATGATCTTCTCGCCCCGACGGGGTGTGCCGGCCGTTTCGACGCGGGCGCTAACGGGACGGTGC  
CCGGCGAAGGCGCGGGGCTGGTAGCGTTGAAGACGCTGAAACGTGCCCTCCAGGACGGCGACA  
CGATCCACGCCGTGATTGCGCGGACCGCGGCAAACAACGATGGTGCCCGCAAGACCGGGTTCA  
CCGCGCCCAGCGCCACGGCCAAGCCGAAGTCATTTCGACGGCGCTGCGCCTGGCCCCGGTGC  
CGGCCGAATCGATCGACTACGTGAGGCCCCACGGAACCGGCACGCGCTAGGCGACCCGATCG  
AGGTAGCCGGCTTGGTGGAGGCCTTCGCCAGCGAGAAGCGCGGCTATTGCCGGCTGGGCTCGG  
TCAAATCCAACCTTGGTCATCTGGAACACTGCTGCCGGCATCGCCGGCCTGATCAAGACCGTGC  
TGGCGCTCGAGCACGCGCACATCCCCAAGTCCTGCCACGTGCCACGCCCAACCCCGCGGCGC

5

82

10

15

20

25

30

35

40

45

50

55

GCCTACACAAGACGCCTTTCCGCATTGCCGCCGACGGGATGGCCTGGCCGCGGCGTATGGCGA  
CGCCGCGGCGGGCGGCGGTGAGTTCGTTCCGCATCGGCGGCACCAACGTCCACGCGATTTTGG  
AGGAGGCGCCGCCCCGCGCGCCCGAGCTGGCGGACGGGCGCAGTCAGGTGTTCTCTCCG  
CCAAGGACGAGGCGGCGCTGGACCGTGCCCTTGCCAACTATGGTGCGGCCTTGGAGAAGCGCG  
5 GCGACCTCGCGGCGGGCGGCGGTGGCCTGGACGCTCCAAAACGGCCGGGCGCATTTCGAATGGC  
15 GAGCCAGCGCGGTGGCATCCGACCTCGACGAATTGGCGGGCGCATTGCGCGGCGAGCGGCCCCG  
GCGCCGTCAAGAAAAACCGAATGGCGCGCGAGGATAAGCCGGTGGCGTTCTTATGTTCCGGGGC  
AGGGGAGCCAGTACCGTGGCATGGGCCACGACCTGTACCGCGAAGAGCCGCGTTTCCGGCACC  
ACCTCGACGCCTGCCTCGCCATCCTCGCCGAACACAAGCCCGAGATCGACTGGCTGGCGTTGC  
20 10 TGGGCTACCGCGACGAGGACGAGCCAACCGACCAGATCGGGACGTCTTCGACAGGGCCCCGAGCC  
GGTCAGCCGCATCGAACCAGCGGAGCTCCTCGACAGCACCGAATTCGCCCAACCTTTGCTTT  
TCTCCATGTCCTACGCGCTCGGTGGCTGTGGCTCGACTGGGGCGTGCGACCCACGGCGATGA  
25 TCGGGCACAGCCTGGGCGAGTACAGTGCTGCATGTATTGCAGATTTCTATGCACTCGATCAGG  
TGCTGCCCTTCATTCTGACCCGCGGTGAGTCATGGCGCAATTGCGGCGCGGCTCGATGTTGG  
15 CCGTCAGCGGTGACAGCGTTCTGATGCGCGAGCTGATCGCCGATGCGCTCGATTTGGCGGCGA  
TCAACGGCGCTGACCAATTTGTCTGGAGCGGGCCGAGCGAGGCTGTCCAAGCCGCGGGGTCC  
30 GACTGCGCGGCGCCGGCCTGCGTGCCACCGAGCTGAACACCTCACACGCGTTCCATTACGCCA  
TGATGGATCCCATTTCTGGAGGAGCTAACGGTTGCCGGTTTCGCGACTTCAGGTCCGTGTGGGA  
CGATTCCGGTCGTTTCATGCGTTACCGGAACCTGGTTGACGGCGAAGCAGCTGGCCGATCCGC  
20 GCTACCACGCGCGTACGCGCGCGAACCAGTGGGTTGCGGGCGGGCCTAGCGACGCTGACAG  
35 GGGAGGAGCCCGCGCTGATGCTCGAAGTGGGGCCGGGCTCGACCCTGGCGGCTTTGGCCCCGCG  
AGCATTTCGAATGCCCGCCTCCCGGTGCTACACAGCCTGCGCCACGCTCGCCAGGCGACGCCCG  
ATCGCCAATACCTGCTCGAAACGCTCGGCTGCCTTTGGCGACACGGGGTTTCCGTGATTGGG  
40 GGGCCCATGCCGGACGTTTCGCGACGCTTGGTTTCGCTGCCCCGCTATCCCTTTTCCGGCGCGG  
25 TGCGCCGCTTAGCCGGCGACCCCTCCGCCTGCTGGCCGGAGCCCGCGCCGTGCGCGCCCCGT  
CGGGAACGCGCCAACCTCAGCGCCGACGCGCGCGACCTCCCGAACACTCCGGAGCCGACATCCG  
GCGCCGTGTGGCGATCAAAGCGCCAATCGCCGCCGCCGATCCCGGCCTCTATCGCCTCTCCT  
45 GCGCCAGGCGGAACGGCGCCGCTCGGTCCGCCCCGATCTCGGTCCGCCCCGCGACTGGATCG  
TCTTCGCCTCTGATTCTCACCTGCTCCAGGCGCTCAGGGCCAATCTCGGGACGCGCGCTCAGC  
30 GGGTGACGCTGGTGACGCCGGGCCAGGAGTACGCAGCCGAGCCGTCCGGGTTTCGGCTGCGGC

5

10

15

20

25

30

35

40

45

50

55

CGGACCAGATCGACGATTACCGCGCCCTGTGGGCGGACTTGGCGCAAACCGGTATTGTGCCAC  
GATACATCGCGTTCCCTCGCCCCGTTTCATGTACCGGGCGCGCATGGCGGGCGATGCCTCGACCC  
TGGACGAAGTGC GCGAGGGCGGCTTCCTGCCCCCTGACCCGCTTGATCCAGACTCGCCCGCCAG  
GCGGACCGAGCGGACTTCTAAGCCTCACGATCGTCACCCCGGCCGCCCTGGCGCTGGGCGACG  
5 AAGCGACGCGCCCGGAATGGGCAATCCTGCACGGGATGGTCGCCGGCTTAAGCCGCGATTATC  
CCGAATGGCGCTTCGTCTCGATCGACGGCGGCGACCCATCCCCGCATCGGTGCGAAGGTCTGG  
CCCCTTGATCGCGCTTCATGCGGTGACGAGGCTGGCCCCGACCCGCTTGGCGCTGCGCGGCC  
TTCACGCTTGGGTTCACAGTGCGAGCACGTTACGCCGGCCACCATCCCTGGGGCGGGTATGT  
GGCGCGAGGGTGGTGTGTACATGATAACGGGCGGATTTCGGCGGGATCGGTCTGGCGCTGGCCC  
20 GCGCCCTGGCTCGAGAAGCTCGCGCCAAGCTGATCCTGGTCGGCCGAAACCTGCCACCGCGC  
CGATCGATCTCGAGGCTTGGGACGCGCCGCGGTTGATTCTCACCGCCGACGTCGCCGACGAAG  
AGGCCATGCGCCGCGTCTTCGATGCCGCGCACGCCGGTTTCGGCGCCATCGACGGCATTCTTC  
ACGCGGCCGGTGTCCCCGGTGGCAGCCTGTTCCGCAACCAATCGGACGCGGCCTTCGAAGACG  
25 TGCTGCACGCCAAGGTTGCGGGTACCCTCGTGCTGCAAGGCCTGAGGGCAATCGATGCGCCGC  
TGTTGCTGATGTCCTCGCTGGACGCCTGGCTTCCCGGTCCCGGTCAGACCGCCTATGCCGCCG  
CCAACGCCTTCCTCGACGCCTTCGCCAGTCTGCGCCGGCGAGAGGGAGAGCCGGTGTACAGCG  
30 TTGGCTGGGACAGTTGGTGCAGAGGTGGGCATGGCTGCTCGGGTCGCTGCCCGATCGGCCGACG  
AACGCGGCCGCTGGCGCGCGAGGGGATCAGCCCTCGCCAGGGTTGGCAGGCTTTGAGCCGGG  
CGCTCGCCCTCGACCCCCCCCCACCTGATGATCTCGCGCACCGACCTGACCTCGCGCTGGCACA  
20 GTCGATCCAGCCCTACGCCGGTCGCTCGAGCGAACCCGAGGTGGCGCTGCCGCGCTGGACCG  
CATCCGCCTGCCAAGCCGTCATCGAGCGTGTTGGTGCGAGCACTTCGCCACCGCCGCGCTGC  
CTCCCGATGGCAACTTTTTTCGAGCTCGGCGCCAGTTCCTTCGACATCGTCCAGCTCAGCGCTC  
GACTTCAACAACAGTTCGGCCGAGATGTCAGCCACACCGTGCTCTACAGTCATCCCACCGTCG  
40 CCTTGCTGGCCGGCTACTTCGCCAATGACCCGACGCCGTCCGGTGCTGCTGCCGACGAACGCG  
25 ACGAAGCGGTGCGTCGCGGCCGCGACCTCTTGAAGAGCCGCGGCGAGGAGTATGACCGTGGA  
GCACGAAACCGGATTCGAAATCGCCGTCATCGGGCTGGCTTGCCGCGTTCCCGGCGCTGCCGA  
CGTGGCCGCTTCTGGCGCAACCTGGTCGAGGCCAAGGAGAGCGTGCGCTTCTTCGAGGACCA  
45 CGAGCTGCGGGCCGCGCGGTGCCCGAGGAGATCTTGCGCCTGCCCAACTACGTGAAGGCCAA  
GCCACTGCTCGCTGATGGCGAAGCTTTCGACGCGGACTTCTTCGGGTTCATCCGCGCGAGGC  
30 CGCCTACCTGGACCCGCAAGTTCGCTCCTGCACGAATGTTGTTGGACCGCGCTGGAGGATGC

5

84

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

CGGCTACGATCCCGCGCAGTACGCCCTACCCGATCGGGTTGTTTCGCGGGCGTCTCCAGCAATCT  
CTCGTTCCTGTTTCGACCGCATCGATCCGCGCGACTCCCCCTGCAGAAGCGCTATGTGGCCGA  
GCTGAACGCGGCCTCCTTCGCCACCCAGATCGCCTACCGGCTCGATCTGAAGGGGCGGCCAT  
TTCGATTCAAACCGCCTGTTTCGACGTCACCTGGTGGCGATTACCTGGCGGCGCAAAGCCTGAT  
CGGCGGCGAGTGCCACATGGCCTTGGCCGGCGGAGCGACCTTGGAGGTCCCCAAAAGCCCCGG  
CTATCTCTACCGCGAAGGCTACATCAACTCGCCGGACGGCCACTGCCGGGCCTTCGACGCCGA  
CGCGGCCGGCACCATCTTCGGCGACGGCGTCCGCATCGTCCTGCTCAAACGCTACCGCGACGC  
CCTACGCGACGGCGATCACGTGTACGAGTGATCAAAGGCTCGGCGATCAACAGTGACGGCCA  
TCGCAAGGTGTCCTACACGGCGCCGGGCAAGAGCGGTCAAGTGGCGGTGATCCGCGCTGCGCT  
GGCGGCGGCCCAGGTAGAGCCGCAAACCATTTCGCTTCGTCGAGGCCACGGGACCGGCACACT  
CGCCGGCGATCCGATCGAGGTAGAGCGTTGACGGAGGTCTTTGCCGAAGCGGGTCGCGGTAC  
CTGCGCCCTGGGTTCGGTGAAGACCAACATCGGCCACTTGGATGTGGCGGCGGGCGTGGCCGG  
TTTCATCAAGGCGGTCTTGGCGCTCGAGCGGCGGTCCTCCCGCCCAGCCTTCACTTCGTCCG  
GCCAACC CGGCCATCGATTTCAACGGGCCCTTCTACGTTTGTGCCAAATCGAGCGGTTGAC  
GGAGAACGGGCGGTTGCGGGCCGGGTGAGTTCCTTTGGCATTGGCGGCACCAATGCCACGT  
GATTCTGGAGGAAGCGCCGGCGCCGGAGGCGAGACTGCCGGCCGGGAGCCCGCCAGGCGCGAG  
TCCGTTCTGTTCCCGCTATCGGCCAAGACGCCGGATGCGCTGGCAGGCCGTTGCCACGACCT  
TGCCGACCACCTGCGGGCGCACCCCGAGCTCCTCTGGCCGATGTGGCCCTCACTCTGCAGAT  
GGGCGGGCGTCTGTCCTACCGCCATGTGGTCCAGGCTGCGACGGCGGAGGAGCTGATTCTG  
CGGTCTGGGAGCGTTCCGACAGGATCCATCCGCAAGAGGCGGAATCGAGTACAAATGGGTGTT  
GGCAGGCGAGGCGATGTGCTTGACCCCGGTTTGGCGCTGTACGCCGATTGGCCGGTCTATCG  
GGAGCGGGTCGACGTCTGTCTGGCGATCGTCGCCAAGCTGCGCCAAATCGACGGCCGGTCATT  
CCTACATGAGTGGATCGAGCGACCGCGGAGGTTCTTGGCCGAATGGTCGACGGCGCTGGCGTT  
CATGTTCCACTGCGCGCTGGCGCAAACCCCTGAGCCAGGCCGGCCTGCACCCGCAGCGCATGTG  
GAGCCGTGGGCTGGGCGGACAGGTGCGCGTGGTTTTGGCCGAATCCCTGTCGTTGGAACAAGC  
GCTGGCGCTGGTGTGTGCCAGACAACGGTTCCCGGCGATGCCACACCTCAGCGCGAACGCTT  
GGTTCGGACACTGGAAGGCTGCCGGTTTCGTCCACCACGATTTTTGATTTCGGCAGACAGCTC  
GGGTGACCCCTGGACCTCGCCGAATTCGCTCATGTGATTTTTGGTGGGTGGCCAAAGCGC  
CTCGCCCAATGAGGCGGAGCTGCGCTCATGGAGCGACGCCGCGCCCGAGCTGGTGACCTTGGC  
GATCGGCCCATCCTTTCTCGAGGCCGCTCCGGGACGGTGGGTCTGGCGATCGACCCCAAGCG

85

ACCGATGACCTGTGTTTCAGCGCACGGTGGCCGCGTTGTGGGAATGGGGATGTGACGTGCGCTG  
GGCTGCGTTCACCTCGTCGACCGGGCGTCGGGTTCCTTCGCTACCTATCCCTTCGTGCGGGT  
AATCCCACGATCGGCGACCCCTTCGCGGAGCAGGCGCGGAGGATGACTTGATTGCGGCGAG  
CGCTTCGCGTCGGCCGGATCGCCGCCCGAGCCGTCGGCAAACCTCGGCAGCGGAACGCCACG  
5 CGCCAGTCAAGCATCGCCTCGGCAACCACACCGGCTCCGTCTCATACTCGGCCAGCGTGGC  
CGTGGCCACCATTCCTCGAAACGTCGCTGCCTATTTGCGGTTCGCCGCCGTGCGTTCACCGA  
CGCCTTCTTCGAATTGGGCGCGTCCTCGCTGGATTGTTGTTCAACCTGGGCCAGCTCCTTTCCGA  
TCGTCTCGGCCGCGAGGTTCCGACCCTGCTCCTCTACGACCACCAACACCGGACCAGTTGGC  
GCTGCGCCTGACATCCGCGGCGCTCAGCGCAGAGGCGCCGCCCTTAAGGGGCGGTTCATCGCGC  
10 ATCGACTTCGCGCACAGCCGCGAGCTCGGCCGCTCCACCGCACCGACGTTCCCGGGGGACGC  
TCACTCGCAGCCCAGCTTCGTTTCGCGAGCAGGACATCGCCATCATCGGGATGGCCTTCGCGG  
ACCGGGCGCCGACGACCTGGACGCGTTCTGGAACAACCTGGTTCGAAGGGGTTCGAGTCGATCAC  
CTTCTTCAGCGAGGACGAGCTGCTGGCGGCGGGCGTCCCCCGGAACATCTGGCCTCGACGCG  
CTACGTGCGGGCCAAGGGGGAAGTACTGGGATGATGGATTTTCGAACCGGAATTTTTTCGGTTA  
15 TTCGGCGCGCGAGGCGGCGGTTCATGGACCCGCGAGTTCCGCGTGTTCACGAATGCTCCTGGCA  
CGCACTGGAGCACGGCGGCTACGATCCGACCCGATGCGCGGCATCGATTGGCGTCTACGCCGG  
CGTGACCAACCACCTGCCTTGGCTGATGCGAACTTTGCCGACCTGACCGAGGAGGAGCAATT  
CGGCGCGCTGCTCCTCACCAGCCGCGAGTTTTTCGCACCGCTGCTCTCTACAAGGTTCGGCCT  
GCGCGGACCCGCTATTTTCGCTGCAAACCGCCTGTTTCGACGTCGTTGGTGGCGATCGGCACGGC  
20 CTGTCGCGAATTGCGCGCGGGTGCCTGTCAGATGGCCCTAGCGGGCGGCGTGACGGCCAGCAT  
CGAGCGCTGCGGCTACTTCCACCAAGAAGGCTACATCCTCTCGCCTGACGGCCACACGCGCAG  
CTTCGACGCGGCGGCGCCGCGCACGGTCTTCGGCGACGGAGTCGGCATGGTGCTGCTGAAGCC  
GCTGGCCCAAGCCTTGGCCGACGCGGACACGATCCACGCGGTGATCAAGGGAATCGGCATCAA  
CAACGACGGCGCGCGCAAGGTTCGGCTTCACCGCACCTAGCCGGGCGGTCAGACCGAGGCGAT  
25 TCGGGCCGCGCTGCGCGACGCGGGGTGGCGTCGAACCGCGTCAGCTACGTGGAGGCGCATGG  
AACCGCGACACAGAATGGGCGACCCGATCGAGGTTCGAGGCCTTGACCCAAGCCTTTCGCGCCGA  
AGCCGACGGTCCGCTTCCGCCCCGCTCCTGCCTACTCGGCTCGGTGAAGTCCAACGTGGGCCA  
CCTGAACGCCGCGGCGGCGTGGCTGGTCTGGTAAAAACCGTGCTGGCGCTCCAACACCGCCG  
CCTGCCGACACGCTGTTCTACAGTCGCCCAATCCACACATCGACTTTGCGGCGAGTCCGTT  
0 CCGCGTGAACGGCCAGACTTCGGATTGGGTTCGCGCCAGAGGGGACGCGGTGCTGGCGGGAGT

55



5

10

15

20

25

30

35

40

45

50

55

GAGTTCGTTTCGGTATCGGGGAACCAACGCCCACCTGATCGTCGAGGAGGCGCCGAAAGCGCT  
ACCGACGACAGCGGCACCTCTGTGACGGAGCCGAATGACCTCGACGCGGGCGACGCCGACGG  
GCTAGTGCTGCCGATCTCGGCCCCGACGCCGACCGCCCTGGCGCACATCGCGACCAACCTCGC  
CAATCACCTGGAACGACATCCGACCATCGCCCTGGCCGACGTCGCCCTGACCCCTCAGCTGGG  
5 CCGTCGCCAATGGCCCCATCGCCACAGCCTGATCTGCCGGAATCGAACGGAGGCGATCAAGCT  
GCTGCGCGCCGTGCTCCACTCCGCGGAGGTGCCGCCAGCTCAGGCGCCGGTCTCGGATGCGCC  
15 GCGCTGTGTTTTTCTTTTTCCCGGCCAGGGCGCCCAATACCCGAGCATGGCCCCGACCTGGT  
TCGAAACTGTCCCGACTTCGCCCTGCACCTGGACCCCTGCCTCGACCAGTTGGCCGAAGTCT  
TCCCGAAGATCCGCGTTGCATCCTGTTTCGGCGATGGCCCCGCCGATCGGCTCGACCAGACGGC  
20 CTACACTCAGCCGCTGCTCTTCTCCGTGTCCTACGCCTTGGCGCGCTGGTTGGGCGATTTCGG  
CATTGCCCCGATGCGATGATCGGCCACAGCCTGGGCGAATACGTGGCGGCCTGCTTGGCCGG  
GCTTTTCTCGCTGAGCGATGCCCTGCTGCTGGTGAGTGAACGCGGCCCGCTGATGGGCTCGGC  
25 CGCGCGCGGAGCGATGCTGGCCGTCCCCTTGCCCGAATGGGAAGTGGAGGAACGCCTGGAGCT  
TCTGGCCGACGACCGAATCAGCATCGCGGCGGTCAACACCGCCGAGAGCTGCGTCATCGCGGG  
15 ACCCAGCGAGGCGATCGAGCGCTGCGCCCAGCGCTGGGCCGCGCAAGGCCTGACCTGTACGCC  
GCTGCGCACGTCCACGCCTTCCACTCCGCGATGATGGAGCCGATTGTGTAACCCCTTCGGCCA  
30 TGTCTTGGCACGGGTACCTTCGCGCCGCCGCGCGCGCGCTGGATCTCGAACCTCGACGGCAA  
GCCGATCGATTCCGCGGCGGTGATGCAGCCGACTATTGGGTGCGCCACCTGCGCCAACCGGT  
CCGCTTTCACGAGGGACTCAGTCACCTGTTGGCCGAGGACACCCATGCTTGGGTGGAAGTGGG  
20 TCCCGGCCGAACCCTGTCTCTCTTCGTCGCCGCCACCCGGCCTACCGTCACCAGCCAATCGT  
35 CAACCCCATGCGCCATGCAGTCGAGTCGACGGGCGACGTGCGCCGGTGGCGCCAAGCGCTGGG  
CGAACTATGGCGGGCCGGCATGCCGGTGCCTGGGAGCGGCAGCGGCGCGGCCGGCATGCCGG  
ACGACGTGTGCCGCTGCCGGGCTACCCCTTCGAGCGGCGGCCCTTCGCGGCCCGAAGACCGGT  
40 GGAGCTGGCGCAGCCCGCGCCCAAGGCGGAGCTGGTGAAAACCCCGATCCCGCGCGGTGGCT  
25 GTACCGCCGCGTCTGGCGCCCTGCCAGGCTGCGGCCGGCGGACTGGCGGTGCAGGCGACCGT  
TCTGGTCTTCGGCGACGGGTCCGAGCTGTGCCGCGCGGCGGTGCTCAGGTGCAGCGCCAGGG  
GCTGAAGTGCGTCTCGATCACCGCGGGCCGCCAATTCGCGCGGGAGAGCGACATGCGCTTCAC  
45 GCTTGACCCCGCTGATCCGCGCCAGCTCGACCAGCTCTTCGCGGCCCTCGATGGCTCAGGCTC  
GCGGCCGCGGTACGTCTGACCTGCTGACCCCTGAACCCGCCCCCGGATGCCTCGGCGATCAT  
30 CGCTCACAGCTACTACAGCCCGATGGCCTTGGCTCATGCCTTGGGCGCCCACGAGATCGCGCC

5

87

10

15

20

25

30

35

40

45

50

55

TGTCTCGATCACCGTCGTACCGCCGGGGTCGTGCGCGTCGCGGACGAAGCGATTTCGCGAGCC  
GCTGCAGGCGCTGATCGTGGGCCCCGTGCCTGGTCATCCCGCAGGAGTTTCCCGGGCTCAGCGT  
TCGGCTGCTGGACGTCAACGTGACGATCCGGCACCGCGTCTGGCGGAGCGGCTCGTGGCCGA  
GCTCTCGGGCACGGATCACATGGTGGCGCTGCGCGGCGGCGAGCGCCTAGTGGCCGATGTCGA  
5 TCAAGTCGATGGCCTCGGTGTGGGGATCGCCAAGGTGCCCTTGCGCCGCGAGGGCCACTACCT  
GATTCTCGGCGGCCTGGGCGATATCGGCTACCACTGTGCCCGCTATCTGGCCCAAACCTACCG  
CGCCAAGCTGACGCTGACCGCGCGTTCGTCACTCCCGCCGCGCGCTCGTGGGAGCGAATGCT  
GCGCGAGGGAAACCTGGATTCCCGGCAGCGCACGCGCATCGAGCGCGTGTGTGCTAGAGGC  
GTGCGGGGGCCGAAGTCCAGACGGCTGCGGTGCACTTGGGCGATCGCCATCGCTTGGCCGATGT  
10 GTTCCGCGAAGCACGGGGCCGATTGCGCGCCATCGCGGGCGTGATTCACTCGGCGGGGATTCC  
GGGACACGTCCACTCGATCGACGAGCTGGTGCCTGCGCGACGAAGCCCAATTCACCGCGAA  
GGTTCGAGGGCTGCACCACCTGGCCGAGGTGCTCGATCCGCTGAACCTCGACTTTTGTCTGCT  
GTTCTCCTCGCTCTCGACCGTCTCGGCGGGCTCGGCTACGGCGCCTATGCAGCGGCCAACGC  
CTACATGGACAGCTTCGCCCCGCCGCCACGATCGGCCGGACGAATGTCGTTGGATCGCGGTCAA  
15 CTGGGACGCCTGGCTGTTTGAAGCCAAGACGTGCTCGGTGCGCGCCGAATTGGCGCGCCTGGC  
GATCGTGCCCCGAGGACGCTCCGGCCCTGTTGCGCGGGTGCTAGAGCGACTTCCGCAATCGTT  
CATCGTGTCCACCGCCGACCTTCGGGCCCCGATCGACACTTGGATCCGGGACAAGAACCGCGT  
CCCCCCCCGAGATCCGAGCGGTTCAACCGCGACCGGACCTGAGCCAGGCGTACGCCCCGCC  
GATCGGCCCCGCTGGAGATTCAACTCTGCGGGCTGGTCTCCGCCTATTGCCGGTTTCGACCGGAT  
20 CGGGCGGGACGATTCTTCTCGAAATCGGCCTCAGCTCGTTGCACTTGATCCAGCTCAGCTC  
GCGCATTCACCGCATCACCGGCAAGGATCTCAATACGACCCAACTGTTTCACTACCCACCGT  
GCGCGCCTTGGCGCTCTTCTCGGCGGCGAACCAGGGGGCTCGCGGCGGAGGAGCCCGCCAT  
GGAGAACCTGTGGCTGCAACGAAGCGATGCGACCTCGATGAGTGAGACCGAGGTGCGCGACT  
GCGGCGCTACCGACCGCGGTGAGGATTTTCCGCGCAGCGATCCGGGACGACTCGCTGAAGAA  
25 GCGCGATAGAAGAACGGAATCGTGTATGAAATACGAAACCACCGGATTGGAATTGGCCGTCAT  
CGGTCTCGCTTGCCGCTTTCCAGGCTCACCCGATCCCGAACAGTTCTGGTCAATCTGCGCGC  
AGGTGCTCCGGAATCCGCCATTTTCAAGCGATGCCGAGCTGAGCCACATCCCCGCATCCCTGCG  
TCACCATCCGCATTACGTCAAGGCCAAAGGCGCGCTGGACCACGCGGATTTTGAACCAGCCTT  
CTTCGGCTACTCGCCCAAAGAGGCCGAGGTGATGGACCCTCAATTCCGGCTGCTCCATGAGTG  
30 CTGCTGGGAGGCGCTGGAGTCAGGCGGCTATGCGCCGAGCCAATTCGCGGGTCCGATCGGCTT



5

88

10

GTTCGCGGCGGCGGCCTTCAACGACGGATGGATCGCCGGTACCCTCGACCGGCTGCGCACCGG  
CGTGGGTTTGAGCTCCCTGGAAACCGCGTTCTTGACCCTGCGCGATTACCTGACCACCCAGAT  
CTCCTATCGGCTCGATCTGCGGGGCCCCAGCCTGCTTGTCCAAACCGCCTGCTCGTCGTCGCT  
GGTGGCGGTCCAGCTCGCCCAGCAGGCGCTGATCTCCGGCGAATGCGCCCTGGCCTTGGCTGG  
5 CGGCGTGTGCGCGACCGATCCGCTGCATTCCGGGATACCTCTATGAACCCGGCAACATCTACGC  
15 GCGCGACGGCGTCTGCCGACCGTTTCGACGAGGCGAGGCGCCGGTACGGTCTTCGGCGACGGGTG  
CGGCATGGTCCTGCTCAAGCGGCTGAGCGACGCCAGCGCGACGGCGATACGATCTGGGCGGT  
CATTCGCGGGGCGGGCGTGAACAACGACGGGCACCACAAGGTTGGCTACACGGCTCCTGGCAC  
GAGGGGCCAGGTGGCTTTGCTTAAAAGTGTTTATCGCGCGAGCCGGGTGACCCGGCGACGCT  
20 CGGCTACCTGGAGGCCCCATGGCACCCGGCACCCGCGCTCGGCGATCCAATCGAGGTCGAGGCGCT  
10 TACCCAGGCCTTCGCCAGCAAACGTCGCGGCACCTGCGGCTTGGGCTCGGTCAAGGGCAACCT  
GGGTCACCTCAACACGGCGGCCCGGCATCGCTGGACTGATCAAGGTGGTGTGCTGGCGCTGAAACA  
25 TCGCGAAGTGCCACCCACCCTCAATCTGCGCCGTCCCAATCCGAAAATCCGCTTCGACGAGAC  
GCCGTTTTTCCCAGTCGTGAGTTGCAACCCTGGCCAAGCGGGACCGGCCCTTGCGAGCCGG  
15 CGTGAGCTCCTTCGGCATCGGCGGTACGAACGCCACGTCATCCTCGAGGAGGCACCGCCGAC  
GGCCAACCCGGCGCCACACGGCAGATTCCGACTGTTGCCGCTTTCGGCCAAGACACCGGCTGC  
30 GCTCGAAGCGAAGCGCCGCGATCTGGCCGGCTTCCTCGAACGCCACCCGGAGACCTCCTTGGC  
CGACCTCGCCTTTACCCTGCAACGCGGCCGCGAGGTCTTCAGTCACCGCGCCTGCCTCGCCGT  
GGAGACCTTAACGTCCGCGCGCACGCGGCTGAGCGGCGAGTCGTGAGCACTTGCGTGGTGGG  
20 CCCC GCGCCAGCGCCATATTTCTGTTCCCTGGTCAAGGCAGCCAGCTCGCCGGGATGGGCCG  
35 CGGTCTGTATCACCATTTCGAGCCGTTCCGCACGGCCGTGATGCCTGTCTGCGCGAGCTGGA  
GCCAGGACTGCGGCAAGCGCTCAGCGCCCATTTGATCCGAATCGCGGCGCGGACCCACCCGA  
TTCGACGACCTTCGTCCAACCCTTGTTGTTCCCTCGTCGAGTACGGGGTGACCGAGTGGCTACG  
40 CTGCTTGGGTGTGCGGCCAACAATGGTGTGTTGGGTACAGCTCTGGCGAGTATGCCGCGAGCCTG  
25 CGTCGCGGGCGTTCTGTGCGCCGTCCGCGGCGGTCTCGCTGCTGGCCGAGCGCGAGCGGCTGCT  
GCGCGACCTGCCAGCCGGCGCCATGCTCGGCGTCCCGCTGGCCGCGGAGGCGCTCGAGGCGAT  
GTTGCCCGACGCTCTCGATCTGGCGGCGATCAACGGCTGTCAGCTTTGCGCCGTGTCCGGGCC  
45 GGTGCGGGCGGTCCACGCCTTCAAGGCCCAACTGGAAGCCCGCGGACATCACGCCCCGCTGTT  
GCACACCGATCGCGCCTTCCACTCGCGGCTGGTAGCACCGGTGCTTGACCGGTTCCAGGCAGC  
30 CGTTCAACACGTGGAGCTGCGGCGGCCGCAAGTACCTTACCTCTCGACCGTCAGCGGGCGATT

50

55

5

89

10

GGAGGCGGATGGGCCGCGAACCCGCACTACTGGGTGCGTCACCTGCGCGACACGGTGCGGTT  
TGGTCCAGCCCTGGAGGCGCTGCCGCCGGTGGATTCCCTTCGTGTGCATCGAGGTGGGACCAGG  
CTCGGCCTTGAGCACCATGGCGCGCGAAACGTTGGGTTCACAGGCGCGACTGATTTTCGTTGCT  
GCCGCGGCCGCGAACGGGGCAAATCGAGCCCGGTCCGGTATTCGAACGACTGGCGGCGCTTTG  
5 GCGCAGCGGGTTGACATTGGATTGGTCTAAATTGACGGGCGGCGAAGAGGGTCATCGAATTCC  
15 CTTGCCAGTCTACCCGTTTCAGCGCAGCCATCTGTGAGCTCCCTGGCGGCGGGCCACACGCC  
TTCGTGCGGCCCTGCAGTCGAATCAGGCGCCATCCTTGCCGAGCGATCCGCAGGGGAAAACGC  
TGAAACCCGGGATTGCCCGCTGCCAACCGCCACGCTCGAGCCCAAGGCGGTGCTCCGGCCCC  
ACTCGAGGCTACCGACGCCCGCAGGTACTCGCGAGCGACTGGCCGAACTTTGGCGCGAGTTGCT  
20 10 AGGGTTGACCTCGATTGGGCCCCGACGACCATTTCCTTCGACCTGGGCGGCCACTCGCTGACCGC  
CACGCGGCTGCGCGCCCTGATTACACAGCGGTTGATGTGATCTCGGGCTCGACGAAATCTT  
CGCTCATTCGCGTCTCTCCCAGCTGGCCGCCCGTATCGAGGCGGCGGCCAAGAGCCGATTTTC  
25 CTCCATTCCCAGCGCGCCGACCAGGACGACTATCCCTTGTCATCCGCCCAGCAGCGGATTCA  
CAGCATCGTCACGAGGGCCGAGGTGCGCACTGCTTATAATTTCCGATCGTCCTCGAGCTGCA  
15 GGGCGCTCTGGATCGAGTGCGATTTCGAGGCGACGTTTCGCGGCATTGTTCCGGCGTCATGAGGG  
GTTCCGCACCCGCTTTGTGATGCGCGATGGCGGGCCGCGCCAGCGCATTTGTACCGGACGTGGC  
30 GTTTCGCCTGCCGCTCACCAGGTGCGAGCCAGAGCAGGTTCCCGGGCGCATCGAGGCCTTCAT  
CCGTCCCTTCGATTTCGAACGCGCGCCGCTGTTCCGCGCGGAGCTGTTGCAGTTGGCCGAGCA  
GCGCCATCTGCTACTTTTCGACATGCACAACCTTAATTGCCGACGGTATCTCGCTCAACCTGTT  
20 20 CGTCGCCGATTTTCGCGGCCCTGTACCATGGTCGTCCGCTGGCGCCGCTGAAACTCCGCTATCG  
35 CGACTATGCCGTTTGGCAAGAGGCGCGGCTGGCCTCCGATGACCTGCGCAGCCAGCGCGAATG  
GTGGCACCGGCGGCTTTCCGCCCGGTCGCCACGCTGGCGCTCCCTCCCGATTTCCCGCGTCC  
GGCGGTGCGCCGCTACAAGGGCCGTAATGTGGTGTTCCACCTGGACCGGGAGATCCGCGACCG  
40 CCTGGTGGCCCTGGCTCGAACCCAGGGGGTCACCATGAACGTGATGATGCTGGCGCTCTGGGC  
25 TGCGCTGCTGCATCGCGAAACCGGCCAATCGGAGCTGGTGGTTCGGATCGCTGCTCGGCGGGCG  
GCCGCACAGCGAGCTGCATCCCGTGATCGGGCTCTTCACCAACTTTTTGCCCTTGCGGTGGC  
GGTCGAGGGATCGACCCGCTTCGATCGCTTCCTTGCCGCTTGCCACCAGGTGTTTCTCGAAGC  
45 CTATCAGCGCCAGGACTATCCGTTCCACTTGTTAGTCCAGGAACTCGTGCCGGTCAGGGACCC  
GTCGCGGTGCGCGCTGTTCCAGACCTCGCTCGTCTACCACAACGAAATTGACGGCAAGACCAA  
30 GCTGGAATTGGAAGGGCTGAAAGTCGAAGTGGTTCCTTCGAAAAGGGTGTGGCGAGGCTGGA

50

55

5

10

15

20

25

30

35

40

45

50

55

TTTGAAGCTGGATGTGACACCTTTTTCCGACCGACTCGAATGTGTTTTGCAATACGACTTGGA  
TCTGTTCTGCGAGGAGACGATGCGCGGCCTGATCGCGCGGTTCCAGGCGTTGGTGGCGGGGCT  
TGTCGCCGATCCGGCGCAATCGCTCGCCGCCGCGAGCGTTTCCGGGAAGCGGGCGCTGCGCGC  
GGGCGTGGCCACGGCAAGCGAATCGTCGCCGCGAGTCACTGCCGCCGCAACCATCGACGGCGTA  
CGCCACTCCCTCACCGCAGTCACCGTCGCCGGTAGTCCTGACGGGACCCGCCGACCTGCCCCG  
GATCTTGGCGGCCTACGTGGGGCAGAACCCCATCCGTTGCGGATCCATCGGGGTCTCATTTTT  
GGAGGCGCCGCTGGGGTTGCGAGCGCTGCGGTGCGCGCTGGACGCAGTGCTCGGAGAACACAC  
CCATTGGCGCAGCGTGCGTGCGGGCGATCGCGCGCGGCGCGTGGATAAGTTGGAATTGACCAG  
CCTGGTGCGGCTCGACGACCTGCGCGGGTTGGTCAATCCTCAGGCGAATGCCTTCACCCTGGC  
TTGGCGCGATCTGGCGATGCCGTTCCGGGGAGGGGCGTCCCCTGTGGCGACTCCGCCTGGCGTG  
GTCGGCTCCATCGCGCTGGTTGCTATTGCTGACGGTTCATCCATTGATCGGCGACAACGGCAC  
GGTCGACCTCTTTCTGGCGGCACTCGCCGATCACCTGCGCCGCGCGTCCGCTTTTCCCGTAGC  
ACCGCTCGATGAGGCCGAGCTGGAGGCGGAGCTGAAGTGGGGAGAGGAAGGGGAGGGCCTCGG  
GCTGACCGCGATCGCGCCGGTCCTGGGCCAATTGCGCGAAAGTCGGCTGAGTCCTGTGGSCCA  
GATGTGGCTGGACGAGGTCTGTGCGCGCCACGACCTCACCCCGCTAGAGGTCTTGGCGGCCCCG  
GCTCCTCGATTGGACACGAAGCCACGGTCACGGGTCGATCGCTTTGTGGACGCCGCTGCCCGA  
GGACCATCCGCTTCGCGATGAAGGCCGCTGCCTCCAGGTTGCGCTGCTGGAGGGGCCCGCGTC  
GCAGCGAGGAGCGGGCGATCCAAGCTGGCTCGAGCAAATCGCCTTGAGACGGGGTACCCCTGC  
AACGGAGGTCGTTTGCCCTACTCCGACCCAACGGGCAGCCATCGACCTCGCGCTGGCCTGGCT  
GCCCGAGCCGCCTCTTCACGGTTTGGTTCGGAACCGTTCAGCCGTGGCCGGAATCTCCATTGGT  
CTGTCCGTTTCCCCTCAATCTCGCGTTCGGGCCAAGCCATCCAATTGCCTACGCGCTCAAGCA  
CGAGGCCACGCTCGCGGTACGGCACGGGCGCGGATCTGATGCGTTTCCTCGACGGCTTGGG  
CCCGGAAAGCTGAAGATTAGCATAAGCGCCCGGCCAAGGGCATCCTAGGATGACGCAAGCCTC  
GGCCGCGTCGACGTCCCAGGTCGCGCCGGAGGTACCCCCGGCCGAAAGGACGACGATGACGA  
TCAAATCCGAGATGTGCGCCGTTGCTCACTCTGCGGAGAGCGGCTTCCGCGCTGGGCCACGCG  
TGGGCGGCGCGATGAAGCGGGGCCGACGCCGGAGCAGGCCGGCGTGAAGCTGCTCCGCGCCC  
CGGTGAAGCGGAAGTGGCTGCCCCCGGCGCCCGTCTGCGCCTGAGCGAGCGGCGTATCCCGG  
AGGTGTGGGCAGGCTACCGCGCGAGCGCGGGATGACCCGAGCCCCGCCCGCGCGGACCAT  
GACGCCGCCCCACGGGGCGAGTCGTCCGGCGCGCCGGCGCGCGTCCGGGGCTTCCGCCGCGGG  
CGGGCAGGTGCAGGATGGTCGGGCATGGTGACGCGTCCGACGTCCGACGGCATCGAGGACGAG

5

10

15

20

25

30

35

40

45

50

55

CTCGCGCCGTTCCCCCGGTCCTGCGCGGCTGGCTCATCGAGGGCGAGCTCGGCCGCGGCGGG  
ATGGGGCGGGTGTTCCGGGCGCGGCACCCGAAGACGCGGGCGGGCGGCGATCAAGGTGCTG  
CTCGGCGACTACGCCCCGCGCGGACGTGGTGGCCCGCTTCCGGCAGGAGGCGATCGCCGTC  
AACATCATCAACCACCCGGGAATCGTCCGCGTCTTCGACTCCGGCGAGCTCGAGGACGGCTCG  
5 CCCTACATCGTGATGGAGTACCTGGACGGCCGGGGGCTGCGCGACTGGGTGCAGGCCGTGCCG  
15 CCCGCGGAGCGGCCGCGGCAGGTCTGCGGCTCGGCTACCAGATCGCCTCGGCCATGGCCGCG  
GCGCACGCGTCCAAGGTCTCCACCGCGATCTGAAGCCGGAGAACATCATGGTGGTTCGAGGAC  
GAGCTCGCGCCCCGGGGCAGCCGCGTCAAGATCCTCGATTTCCGGCATCGCGAAGGTCTCTGG  
GGAGGTCTGCCCCGAGGTGCTGGAGCTCGAGGGGAGAGGCTCCCTCGCGCCCGCGTCCGCGTCC  
20 10 ACGATCCGCACCGAGCTCTCGACGCGGCCGCGCCGACGGTGGGCGCCACGACCGGCCAGAG  
AGCCCCGTGGGCGCGAGCGCCACGCCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCAGAGAGC  
GCCCTGGGCGCGAGCGCCACGCCAGAGAGCGAGGCCACGAGGAAGACGCGCTCCGGAGCCTC  
25 CCCGTCTGACCAGCGGCAGGCCCGCGATCCACCCCGCGCCGTCGAGATCCCCCGCGAGGCG  
GTCTCCTCCGCGGCGTCTCGCGGGTCTCGCGCGCTCGATCGAGCCAGGCGCGCCCGCGCCGCG  
15 AGCGAGGGCGCGGGACAGCCACGATGCCGTTACGCAAGAGGGCGTGTGGGGCCTCGGGACG  
AGGAGCTACATGGCGCCGAGCAGGAGCGCCACTCCGGGAGCGTGGACGTGAAGGCGGATGTC  
30 TACTCGCTCGGCGTCATCCTCTATGAGCTGCTCGAGGGGCGGACGCCCGACGCGCCGAGCGCC  
GCGTGGCCGCCCCCGATGAGCGCCGCCACGCCGCCGATCTCGTCCGCCCTCGTCCACCGGGTT  
CTGGCGTTCGATCCCGATGCGCGGCCGCGCATGGCGGAGGTGGCGAGCGCGCTTACCGGCTC  
20 GGCCGGGCGAAGAAGGAGCTCGACGAGGCGCTCTCGAGGTGGGTCTCGGCGGAGGGGCGCCG  
35 GGGCTCTTGCCGTGCGGCTATGCTCTTCTCGAACTGGTCCTCCTGGGCCCTGGGAACCTTATAC  
GATTCTTTCCAGCCTGTAAGTGCATTTTTCTTTCAATATCGTCCTCTCTTCATATACGAGGTG  
AGTTCTCTGAGGTCTCTATAAGTCTGGGGTGTCTATTTCGGCCTCTTACTTGTTACTTCGC  
40 CTTCTTAGGAGTTTTTCTTAATTTGCCCTCTTACATTCCCGTATTCAATTCTAACTGGGCCC  
25 TATCTCATTTCGCTAATACGTTTCTGTATTGTGTACATCTCCTATCATGTGTCAATACTTGTTT  
CTGTTTATCATTATTCTTATTGTTTACGCTCTTATTTCAATCATAGTATAACATTAGTTTACT  
GATTATCGCACTTGAATTCGCG

or its complementary strand,

30

5

92

10

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

15

5 (c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

20

10 (d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

25

15 8. DNA sequence according to claim 6 selected from the following

30

(a) open reading frames:

35

40

45

50

55

<u>Nucleotide Position</u>				
20	ORF1	1666	- 1	Seq ID No 3
	ORF2	1605	- 3338	Seq ID No 4
	ORF3	6100	- 3398	Seq ID No 5
	ORF4	7110	- 6374	Seq ID No 6
	ORF5	9590	- 8433	Seq ID No 7
25	ORF6	11393	- 9855	Seq ID No 8
	ORF7	13656	- 12712	Seq ID No 9
	ORF8	15374	- 18984	Seq ID No 10
	ORF9	20003	- 27889	Seq ID No 11
	ORF10	28251	- 29402	Seq ID No 12
30	ORF11	31720	- 30401	Seq ID No 13

5

93

10

ORF12	31982	-	32932	Seq ID No 14
ORF13	33128	-	33613	Seq ID No 15
ORF14	33661	-	34007	Seq ID No 16
ORF15	35611	-	35255	Seq ID No 17
5 ORF16	37856	-	35730	Seq ID No 18

15

or DNA sequences complementary to said open reading frames,

20

(b) DNA-sequences which hybridise under stringent conditions  
10 to regions of DNA sequences according to (a) encoding proteins  
or to fragments of said DNA sequences,

25

(c) DNA-sequences which hybridise to the DNA-sequences accord-  
ing to (a) and (b) because of a degeneration of the genetic  
15 code,

30

(d) allele variations and mutants resulting by substitution,  
insertion or deletion of nucleotides or inversion of nucleotide  
segments of DNA-sequences according to (a) to (c), wherein the  
20 variations and mutants offer isofunctional expression products,

35

and peptide sequences corresponding to said open reading frames

40

SEQ ID No 19 (>ORF1)  
25 VDPEREAVTLGLAFNRAQGRTYARGPEARAEYIGTAMRAADVIEDRFEIERLAVSGGMGDVYR  
ARDRVSGQAVALKVLQGASANDLRRFAEAEALVTLRLPGVVQYVAHGVGTGAGRPYLAMEWLD  
GVTLEERLAGAPLTLAESVALAARVATTLGAIHWLGVVHRDLKPSNMLVGGAVERTLLDFG  
45 IARHLRLAPTLTSPGAVLGTPGYIAPEQVRGDAPVDARDVFALGCVLFQCLAGRPPFLGNSAL  
ALLMRVVL EEPRLGELRDGIPEPLERLVARMLAKNAGERPRDGAAAAAELAAVAGEGLSIGA  
30 SAVAAPAAPGEAITTAERKVMCVILAEDGGAEAGATLSEDDGAARAEALRDIAARHGGRLDRL

50

55

5

94

10

QARWWLVALSGAESPTDLATRAAHCALALRAALGGVPVSVATGLAEVEARLPVGELVDRVAQL  
IAGRDGLSPPEIRLDDATASLLASRFETVQGPGGCWL RGPKEEPDAVPRLLGKPTPCVGRERE  
LSQLATEWRHCVDEPSANAVVVVGAPGLGKSRLAWEFLRTLEQREGAAI

15

5 SEQ ID No 20 (&gt;ORF2)

20

VRPCARLNASPSVTASRSGSTAAGSVHASTSACVEQPATGRTQPASPRWPPGAAALRLTSAMP  
RWFNTAGPCNPADHYMLPAEERLPAVRDLVDRKAYFVLHAPRQIGKTTSLRTLAQDLTAEGRY  
VAVLVSAEVGAPFSDDPGAAELAMLAEWRTAGAQLPADLRPPFPDAPAGQRI GAALRAWAQ  
AAPRPLVVFLDEADALRDATLVSLLRQIRSGYPDRPRDFPHALALVGLRDVRDYKVASVDSGR  
10 LGTSSPFNIKVESLTLRNFTTRDEVATLYAQHTAETGQVFRPDAVDRAFELTQGQPWLANALAR  
QLVEVLVKDRAQPITSANVDRAKEILIERQDTHLDSLVDRLREPRI RAVIEPMLAGTALPSVP  
PDDLRF AIDLGLVRMTAEGGLDVANPIYREIIVRELAFPIRASLPQIKATWLTQDGRLDADRL  
LDAFLSFWRQHGEPLLGAAPYHEIAPHLVVMAFLHRVVNGGGTVEREY AIGRGRMDLCVRYAG  
25 ETLAIELKVWRDGRPD PVAEGLAQLDEYLAGLGLDRGWLILFDQ RSGQPPIAERTRRERALS P  
15 AGREVAVIRA

30

SEQ ID No 21 (&gt;ORF3)

35

VTIKKTFRSIDPATLPKHFDSPVAELRLADLWEADGTYRYDPSRPREETFVVDTPPPTASGSL  
HIGHVFSYHTD VVRQRRMRGFNI FYPMGWDDNGLPTERRVQNYFHVRTDVRTPYERGLTLP  
20 QAAPETIKKEPPRIVSRPNFIELCHKV TREDEQVFKALFRRVGLSVDWRNEYATIDDHCRRTA  
QLSFLDLHEKGHLYSVFAPTMWDVDFQTAVAQAEVEDRPQSGAFHDIAFAVEGTAEELVIATT  
RPELLAACVGVT AHPEDPRYQHLFGKTALTPIFRAPVPIFPSPLVDREKGTGILMVCTFGDAT  
DVIWWREQKLPLRQMLGKNRVL PVTFGEGAWESRDPAAANAAYAPLQGRGVKQARA AAVVELL  
40 RREEHAAAPGRGPALRGEPRPIERAVKFYERGDQPLEFVPTRQWFVRLADKKAELLEYGDKIK  
25 WHPDFMRLRYRNWTEGLQGDWCISRQRYFGVQFPVWYPLDAEGNPDHSRPLLATREMLPVDPT  
VDVPPGYEASQRDQPGGFTAESDVFDTWFTSSLTPQISSHWGDDPARHARLFPADLRPQA HDI  
IRTWAFYTIKAMLHESSVPWHHVAISGWILDPDRKKMSKSKGNVVT PMHLLDTYSSDAVRYW  
45 SASARLGTDTAFDEKVLKIGKRLVTKIWNASKYVLSQSAEVHPISEELDRALLHKLSAVVDDA  
TRSFDEHEFAAALERTEDFFWRWFTDAYLELAKARARGE GAGEAARGSAVAALRLGLSVLLR  
30 LFAPVLPYITDEVWRWVYAEETGDTSIHRAKWPSAADFAAVAAPSDPGLLDLAAAAMA AVNKR

50

55

5

10

KSELGASVGRVVTDLALGANAATLARLKPALGDVLTAVRAGAHALVRPELADGEVLVVRCELE  
PAAAAAAGAGGAAASEE

15

SEQ ID No 22 (>ORF4)

5 MIHAEPFEARLVAARPLSPFVRELSFERADGRSFLFEAGQWVNLVLPPLPGGEVKRAYSIASAP  
DGSPRFDLAVTLVQGGAGSEHLHRLEPGATLRAIGPHGLFTRDPGDSAPSLFVATGTGITPLR  
SMLRASLRAGLAAPHLWILFGARFEEDVIYRDELEALARGSDRIRYEITLSRGGPSWAGRRGY  
VQAHVPELYRELAEKSGDPAPHVFICGLDRMVSSVRELARGE LGVHRKHVHVERYD

20

10 SEQ ID No 23 (>ORF5)

MKSLPSDRAARLAQSDIRMTLACAKVHGINMSQGVCDTPVPSVILQAVKEAMDRGCNTYSRF  
DGIVELRHAI AAKLARHNGIAADPETDITVSAGATGAFQATCMALLNPGDEVLLFEPFYAYHA  
25 QAILAVEAVPRYVTARSLSWNVDGDELERAITPKTKAIVVNSPGNPSGKVFGRMELEQIADLA  
CHHDLMVITDEIYEYFIFDGREHVSVASLPRMSERTITIGGYSKTFSITGWRIGYSVADARWA  
15 KAIGAMSDLLYVCAPTPLQHGVAAGIRGLPRS FYTG LAQGYERKRDRFCRALEKAGLPPCVPQ  
GTY YVLADVSR L PGRTGRERAIYLLDETGVAGVPGDAFFEGTQGSRFMRFCFAKTDEDLEEAC  
30 QRIEQLA

30

SEQ ID No 24 (>ORF6)

35

20 VSDPRKERLGDM DLEEFRRIGMRIIDWAADYLGHPDRYPVFPAIRPGDVKGRLAPTPPVEPEP  
MDAVLTD FEQIILPGITHWNHPRFFAYFANTASGPGILGELLAACLN VNVMLWRTSPAATELE  
ELVLSWLRQMLDL DAGLHGAIMDTASTASMVAIAAARDSAEPTIRLRGMAGQRRMRLYASEQA  
HSSIEKAAITLGIGQEGVRKIPTDPAFRMVPEALRAAVVEDLGAGLRPFCVAATVGTSTTSV  
40 DPIPAIVSVCREHGLWLHVDAAYAGMAAIVPEHRDVLACGEGADSLV VNP HKWLF TPMDCSVL  
25 YVRDADRLKRAFSLVPEYL RTEGDVTNYMDWGIQLGRRFRALKLWMIVRYFGHEGLAARIREH  
LRLGQQLAQWVDADPDWERLAPTPFSTVCFMRPSALACIMRSADAEARESIERELDRLNEAL  
LDEVNKSGRVELSHTRLHGRYTIRVAIGNIRSDEVAVREAWEC LRAAGARLCADERFVSCSRS  
45 ADEGRGKS

45

50

55



5

96

SEQ ID No 25 (&gt;ORF7)

10

MRREEPVLEAFYERYCAAPRETSYHVELPVDVELHQEAAPALPQARSLELAGRVALVTGSSRG  
IGKAIALRLAEQGADVAVNYHSNKDAAEQTAAEIRALGRRTMVVQADVTRPNAAAELFSSVEA  
QLGPIDILVNNVGDDFFKPLAAMTDDEWRNVMDSNLSSVHYLCRAAVARMRQRKSGRIINIGL  
5 SPTYAIRGAPNVAAYSIAKTGVILITRSLATEEAPHGILVNCVSPGLIDNGYLPPAQKEWMER  
15 RVPMGRLGRASEVADAVAFLASDRASYVSGANIAVAGGWDWTDGRGTEHDDRVDLFIGHEEP

SEQ ID No 26 (&gt;ORF8)

20

10

MSGRFPGARNVEELWQKL RAGVECVVTFTEAEALAAGVSREMLANPSYVRRGAPLDGVELFDA  
SFFGFSPREAESMDPQQRIFLEVAWEALERAGYDPDAHSGPIGVFAGSAPSGYHSLAQSDPEI  
LGALGHYQLTLNNDKDYLTTHASYKLNLRGPSVCVQTSCTSLVAVVMACQSLNHECDMALA  
GGVGIHAHQRRGYLYQENGISSPDGHCRAFDVAAGKTVGGSGIGIVVLKRLADALADGDHVHA  
25 VIRGAAINNDGSSKIGYTAPSVQGQAEVIGMAQALAGVEPDDISYIEAHGTGTPLGDPIEIAA  
LTRVFRAKTARRQFCAIGSLKTNLGHLDAAAGVASLIKTVMALHRELPPSLHFERPNPKLEL  
15 ESSPFYVNTRLTPWHAARGPRRAGVSSFGIGGTNAHVLEEAPAPPPSGPSRRWQLLTLAARS  
EAGLARATADMIEHLDRHSGTSIADVITYTSHVGRRAWPFRRVVGESAADLRAALASEGSPRS  
30 ISSCQAARERPVVFLFPGQGAQHLFMARELYEVEPIFRQSLDRCAELLRGPLGLDLRQVLYPA  
EGORDDAEQELGRTAIAQPALFAIELSLAKLWMAWGIVPQAMIGHSVGEFAAACLAGIFREED  
ALRLVAERGRLMQMPPGAMLAVPLAEPELAPYLSDDISLAAINGPALSVVAGPIEAIDALAA  
20 ELLDHGLSCRRRLHTRHAFHSKMMAPVVDAFTRCVSAVERRPPSGHFLSTLTGGWISPEAATIP  
35 AYWARQLVEPVRFQAQAVRQLLSESTWLWLELGPQTLSPLVRQQARADGGQVVVASLPRAKDA  
GADHLAVIEALGRVWSAGGTVDWKRFEHEGEARRRVLLPTYPPFERQRYWASPRHTSAPPEAIK  
PLLAKNPNVADWFFLPARRSDPPVSFDAQAVTTRRSTWLFIGDEGLGAALVEGLARRGHEV  
40 VAVVTGERFEQTGTQRYTIDPAANGDVASLFARLEIEGRMPDRIVHAFCTSPADGARIERGAA  
25 LEIERRLGFDSELLLLAQVIAAQRHHPKPLMLGVITTRAHSVIGTEIEPLRALVLGPCRVIPOE  
IPHVSCRNIDIDLPGEGGRAEIAARLIADLERESPDSVVAYRGGRRWVESIELTDVGRRSAGA  
APRLRQRGAYLITGGLGGIGLVAAELLAREAHARLILVGRGTGLPARQGWDDWLAHAGDATS  
45 RKILRIRALEEAGAEVKIAAADVSDFNAMRSVIEEARTFRGRIDGVIHSAGIASGGMIQLRTP  
MAAWRVMA PKVGGTLVLDALLRDERPDFLLICSSSLASLVGGATQIDYCAANAFDAYAQSREG  
30 EEGCRVISVQWDTWSDVGMAVDFKLPADLQEGRRRESLKRGISSEGAEVLGRILSAGMSGPLA

50

55

5

10

ICTSDLPAYKQSVTTTRRSQHEQTPAARPMHSRPTTTGAYVAPETETERRIAIIWQDLLGLEQV  
GANDDFLQLGGHSLLATQVLSRVLQTLKVGISLPQFFDAPTVAGLSRLVDAARAEGAGPVAPA  
IGRVERDAYRIKPPAAEQAARTKP

5 SEQ ID No 27 (>ORF9)

15

20

10 MEPVGGVDMNQPAKQQETCVFPTSFAQRRLWFLDQLEPGSAVYNMPASFRTTRGPYDVDSLVR  
VNEIVRRHESLRTTVDDIDGEPVQVIAPSLRIEVPVVDLSEIDEPEREAEARRLMAEESRRPF  
DLTRGPLLRAKLLRLGEADHVLILTMHHIVSDGWSMDVLFKELSTLYAAFHEGRPSPLPELPI  
QYADFAVWQRELLQGEVLESHLGYWREHLRGAPTLLLELPMDRPRPPAQTFRGSQRAFRPLPSL  
20 QQAVQALSROEGATPFMTLLTAFSVLLSRYARQSDLVVGTPIANRTRAELEGLIGFFVNMLAL  
RIDLGGDPSFRELLGRVREVTLGAYAHQDLPPERLVEELSPGRSPSHSPLFQVSFTLQNTMPD  
ATNRADIASGGAPLVEMKAAKFDLILELSESPQGLLGTFEYNTDLFDAGTIERMAGHLEVLLS  
25 SAVAAPDRPIAELPLMGAEERSRVLVEWNSTAALYPEDHCMHELFEQQVERSPEATAVLLQQQ  
TLTYRELMNRANQLAHLRLSLGVGPEVRVGLYLSIETVVAILGVLKAGGAYVPLDPTYPSE  
15 RLGLMMADAAPSVLLTQASLLSKLPPHGDATLVQLDALHEALSRLPHHTPRSGVTAQNLAYVM  
YTSGSTGRPKGVLEHRGLCNLPTVQAKLYGIAPGDRLLQFAPLCFDTSFCEIALALLSGATL  
30 VMGTADELLPGPPLVELLKKHAVTAMLLAPTIVLAALPEQQAALPLRVLTIMAGEACPAELVKR  
WKAPGRRLFNSYGPTETTIWASSAADLSDERIPPIGRPIANTQIYVLDEALEPVPVIGVPGEIF  
IGGVGVARGYHGRPDLTAERFVPDPFGQTKGARLYRTGDRARWLPDGNLEFLGRNDEQVKVRG  
20 VRIELEEIIRAALLKHPAVAQAVAVVREDTPGDKRLVAYVVGRRGARVTAELRQSVSERLPAT  
35 MVPSSFVALDALPLTPNGKVDRRALPEPEQSAGGEDHVAPRNAVEEELARIWASVLRRLERVGV  
HDNFFEIGGDSILSIQIVVRAQQAGLRLTPRQMFQHQTI AELSTVARAVEAVHVEQDPVTGPA  
PLTPVQRWWLEQEAAPHFNQSI FLEVRERLDESALEQAI AHLIDHHDALRLRLARDERGAH  
40 QVFAAPGGSTPFQRVLDLALPSAEQISAMEKAASEAQASDLAAGPVVRAVLFDLGEVAPQRL  
25 LVIAHHIAVDSVSWRILLDDLF GAYEQARRGEAVRLPPKTTSVKRWAELLTEHAGSEAVKAEL  
GYWLDSSRRTVAPLPVDRRAGEDVWGSARHIVVSLTPEQTEQLLREVPQAYRTRIDDALLTAF  
AQAIARWTGSPAVLLDLEGHGREELAGVDLTRTVGWFTAMYPILLRVDAADPGEALKSIKEQL  
45 RAVPGRGLGYGLLRYLRSDTIAEVRALPQAE LCFNYLGQLDQAIPEAAPFRPAREYQGSERSP  
GAHRAHLIEVNASIANGRLYATWTYSERRHEPETIERVAASFVTALRALIAHCTLPEVGGNTP  
30 SDFDKVRLRQETIDALDAIDAGPGPSARGSRIEDVYPLSPLQEGILFHTLYATDYTAYVEQFH

50

55

5

10

15

20

25

30

WTLEGDFDAEAFTRALQDVVARHAALRTSFAWERLDAPLQIVRTGAVLPVEHQDLRGLAAEEQ  
TAHISRYVEAERQRRFDLRKAPLMRAGLLRLRKDAWCLVETIHHLILDGWSTQILLKEVFTLY  
EAHRGHRGHLALELEQPRPYGDYIGWLAKQDQVRTAAFWRRELEGFSAPTPLGVDRAPVPHDDG  
GPRFGWRRIALSGDDAARLAARQHQLTMTSTLVQGAWALLLSRYSGDPDVLFGMTVSGRSAP  
5 IPGIERMTGLFINTIPVRVREPADASVLAWLKALQEHEAEELLEHEHSPLVEVQAHSDVPRGTP  
LFESLVVFENYPVQVIFEAPPVEGPTRAEEGLRMIDAQYISDPYPLTVVAAFHGTLYLNIGY  
ERRRFDDQAVERMIGHVTLLRGFVQRPETSVRDLPLLTAAEEERTQLHAWNATAAPYPEGHCM  
HELFEQQVERSPEATAVLLQQQTLTYRELNIRANQLAHLRLSLGVGPEVRVGLCLERSIETVV  
AILGVLKAGGVYVPLDPTYPSERLGLMMEDAAPSVLLTQTSLLSKLPPHGDATLVQLDALHEA  
10 LSRLPHHTPRSGVTAQNLAYVMYTSGSTGRPKGVLEHRGLCNLPTVQAKLYAIAPSDRLLQF  
APLCFDTSFCEIALALLSGATLVMGTADELLPGPPLVELLKKHAVTAMLLAPSVLAALPEQQS  
AALPLRVLAMAGEACPAELVKRWKAPGRRLFNSYGPTETTIWASSAADLSDERIPPIGRPIAN  
TQIYVLDEALEPVPPIGVPGEIFIGGVGVARGYHGRPDLTAERFVPDPFGQTKGARLYRTGDRA  
RWLPDGNLEFLGRNDEQVKVRGIRIELEEIRAALLKHPAVAQAVAVVREDAPGDKRLVAYVVG  
15 RGGARLTAAELRQSVSERLPATMVPSSFVALDALPLTPNGKVDRRALPEPERSAGGEDHVAPR  
NAIEEELTRIWADVLGAKRVGVHDNFFDLGGHSLLLVRVHDLRGQRFDRPPSMVDLFTYPTVA  
SLARFLGERANGKQSPREAAADVTERGRRRLEARARRAKAIRGPT

35

40

45

SEQ ID No 28 (>ORF10)  
20 MKHNIGWLLPAALATLAFVPACSPNHGEDAPSVTSAESGAAPSADCVALGAKLQAALDGAAAA  
QKAPGAAAQVSGDCVWRGATGVSDLVASTPTKPGDLFRIGSITKTFVSTLILMLRAEGRSL  
DDAVSKYVKGIPAGDQMTLRQILGHTSGLFDYTYSPALGQMI EVDPTRAFAPAE LIALATAEA  
PYFAPGAGFRYSNTNYIVAGLVAEAVSGGTLAGLLRTRILD PVGLAHTYLDGAEPVQGLIRG  
YGDYAGGLVDITDQLSPTEAWAAGALVSNVDDLNRFFALLISHELLSSDELQDMTTWTPTMWP  
25 HEPGYGLGLIERDSALGSLNGHCGIIWGFQSASYGVPGRGDAITALINRSDGDAARLVDELAK  
VVKER.

50

55

SEQ ID No 29 (>ORF11)  
MSIDRAVLEQLDRVGGRLAEGKALKLLEDIAWPREVEERFFAAGEDRLPEVEYRVDRDGLARR  
30 VAEARELLGAIDGDAPALGWLRDNVRAQIQAAELLEAAGTRAFSARSQELYGGARSRFFGGSL

5

10

RNIDLAEHLTERLRVHGWDEASDPEEEPLDAGALRDMLAARVAGRAPRLDLEITVDPRVTAKV  
VAGMSRVRIRPEATFAAWEAEGLWHHEVETHALTAHNGAAQPRCAFLRSGGPRTTTRTQEGLAI  
FAELYSRSLSIGRLTRLAERVRLVDMAEQASFLDLYRHLRERGAERRDAYFDAQRVCRGGLV  
EGGAPFTKDACYLAGLLEVYAFLLAAVLRGGLRDEVELLVCGRIALDDIAVLAE LRAAGVLERP  
5 RYLPGLRAWQTLTPYFAFTSFMDGIDLGPVERHFQELLRVAADARPAGEGRRRRGRPREG

15

SEQ ID No 30 (&gt;ORF12)

20

MSESVAQLEEHRAALTGH CYRMLGSVVDADDAVQETMVRARSLDKFDGRSSLRTWLYRIATN  
VCIDLRADRARRARPIEEGPVGTVDDALETRPRTHWLEPVPDAHALPADIDAAERAMLRQSIR  
10 LAFVAALQHLPPKQRAALLLTEVLGWSAAEVADSLNTSVAAINSALQARATLASRDLGDARP  
SLPEPQSALLDRYVNAFERYDVDALTALLHQDATLSMPFFTLWLRGHESIRAWLVGPGAGCRG  
SRLIPTAASGSPAFAQYRPAPEGGHRAWALIVLDVAGDRIVSMTSFLDTETLFPRFGLPLDLP  
A

25

15 SEQ ID No 31 (&gt;ORF13)

30

VTIASIDHRDQDLMTGPQAKAPARAAAPDAAPSRRVWAGRVLSGLATLFLTFDAAVKVLKLF  
PAEASTAELGFPAHLVPTLGYLQIACLVAYLI PR TAVLGAILWTGYLGGAIAIHVRVENPLFS  
HTLFPIYVAFLWAGLWLRDRRVRALTASPSSQGR

35

20 SEQ ID No 32 (&gt;ORF14)

MTTKNPRKLFVNLSVRDLKRSMEFFSKLGFEFNPQFTDEKAACMVVSEEAYVMLLVESFFKTF  
MKKEICSTSTHTEGLFALSCSSRAEVDDMVKKAVAAGGSHAMPQDHGFMYGWSFYDVGHHW  
EVMWMDPKAIQP

40

25 SEQ ID No 33 (&gt;ORF15)

MTPSERLDATFAALADPTRRAILARLASGEASVTELAKPFAMSQPAISKHLKVLERAGLISRG  
RDAQRRPCRIEAKPLEDASGWLDNYRRFWEGSYERLDDLLEELKERESKGERSKR

45

50

55

5

100

SEQ ID No 34 (&gt;ORF16)

10

VAPASAPAAGGRDAAPFLDEAAQWLRGEQAPASRPAGEGPAGRLPGRVLVADDNADMREYALR  
LLVAEGWTVEAVADGRAALERARAHPPDLVLTDMMPRLDGFGLLRALRADDRTRGVAVVMLS  
ARAGEEARVDSLEAGADDFLVKPFSAKELLARVRIHVELARRRREAEGQRQYLNDLFMQAPGP

5

IAILRGPEHVFEVVNPLYQRLVGGRSLVGEPIRAALPELEGQGIWELLDVVRTGEPVIGKEL

15

PVRLDRRGDGTTEEVEFFNFVYQPMRDRDGAVEGVFVFAFDVTDQVRARRRVEALVEALKLADQ  
RKDEFLAMLAHELNRNPMASISLSLTLLDDADGDGPASARYREIARRQMGLVRLVDDLDDVSR  
ITRGTVELRLEDVDLAAVVQSAAAARPAVEARRHDVSLSVGPGDFGMRADATRLEQVVTNLL

20

10

GLGLGLTLVRRLLLELHGGSVAAASAGPGQGSEFTVRLPLGPGAAPQPAPSAGPPPPREGPPPA  
QRDEPPPPPAQRAEAEAAADRRRLVVEDAEDVRRVMRAYIEALGHEVTVAVDGLEGVKKLL  
ELRPEVAFVDIGLPGIDGYEVARRARAAPGGEALYLVALSGYGGPDDQARSRRAGFDLHLTKP  
VVGATLQDVLTPRT

25

15

9. DNA sequence according to claim 7 selected from the fol-  
lowing

30

(a) open reading frames, and peptide sequences corresponding to  
said open reading frames:

20

35

pEPOcos6\_ORF1 sequences:

(1) nucleotide sequence

40

Seq ID No 35 (&gt;pEPOcos6\_ORF1.seq)

25

GGATCACCTGCGGCGCGATCGCCGACCTCGTGCTGGTGTTCGGCTCGCTGGATGAGAAGCCGG  
CGGCGCTACTGATAGAGACGGCGACGCCCGGGCTGCGGGTGGAGCGGTGCGGGAGATGCTCG  
GCTTTCGGGCGGCCCCACCTGGCGAAGCTGTCTTCGACGGTTGCGAGGTCCCCGAGGCTCAGC  
TGATTGGCCGGCCCCGGCTTTGCGCTGATGTATCTGGCCCCCTACGCCCTGGATTTCGGTCGGG  
TCAGCGTCGCCTGGGCCTGCCTGGGCATGATCCGCGCTTGCTGGAGACCTGCGCACAGCACA  
TCCTCACCCGCCGCACCTTCGGCCACCTGCTAGCCGATCACGGCATGATCCAAACCCTGATCA

50

55

5

10

15

CCAACCTGGGGATTACCAACCAGGCGACGCTGCTCCACACGCTGCAGGCCTGCCGCGCCAGGG  
ATCGCGGCGACGTGACCGCCTCCGAGGCCACCCTCGCCGCCAAATACCTCGCGTCGCGGACGG  
CGGTCCAGGAGACGACCAACGCGGTCCAGATCATGGGCGCGCTGGGCTGCGACGAGGAGGGCG  
CGATCGCCCGCCACTTCCGCGACGCCAAGACGACCGAAATCATCGAAGGCAGCAACCAGATCA  
5 TCGAGGCGCTGCTGGCCAAGAACATCGCCCGCGCCGGTCGCGACAACATCGCCGCTTCCTCG  
15 ATGCGGAAGTCGAGCCCGGTCTGGGCGCGAGGCGCACCA

(2) peptide sequence

Seq ID No 36 (>pEP0cos6\_ORF1.pep)

20

10 ITCGAIADLVLVFGSLDEKPAALLIETATPGLRVERLREMLGFRAAHLAKLSFDGCEVPEAQL  
IGRPGFALMYLAPYALDFGRVSVAWACLG MIRACLETC AQHILTRRTFGHLLADHGMIQTLIT  
NLGIHHQATLLHTLQACRARDRGDVTASEATLAAKYLASRTAVQETTNAVQIMGALGCDEEGA  
25 IARHFRDAKTTEIIIEGSNQII EALLAKNIARAGRDNYRRFLDAEVEPGRAGGAP\*

25

15 pEP0cos6\_ORF2 sequences:

30

(1) nucleotide sequence

Seq ID No 37 (>pEP0cos6\_ORF2.seq)

35

40

45

ATGACGAGCGCGGTCCCGACGCGTCAAACCAGCCTGCTCGACGACTTCGAGCGCGTCGCCGAC  
20 GTCGATCCAGAGCGGATCGCCGTCCACGCGAGCGAGACGAGCCTGCGCTATGGCGACATGAAT  
GCGCGCGCCAACCGCATTGCCCACGGGCTACGGGCGCGCGGGATCGGGCCCAATCAAATCGTG  
GCGGTGGCGATGGCCCGCACGCCCCGAGCTGATGATCGTGCTGTACGGCATCCTCAAGGCCGGC  
GCGGCCTACATGCCCATCGCCCGCGACGCGCCGCGCTGCGCCGCGATCATATGCTGCGCGAG  
AGCCAGGCTGCTCTGATGATCGCCGACGAAGAGATCGCGGGACTCGCGGCCCGGGTGCTGACG  
25 CCGGCCGACCCGTTCTTCGCGGCCATGCCGGACCACAACCCCGAGCCGCGTCACGACCCGACC  
GACCTGATTTACGTCATCTACACCTCGGGCTCGACCGGCCAGCCCAAGGGCGTGGCCATGGAG  
CACCGCGCCGTGTGGAATCGCCTGACTTGATGCAGGCCAGTATCCAATCGACACGCAGGAC  
GTGATCCTCCAAAAGACGCCGATCCTCTTCGACGTGTCGGTCTGGGAGCTGTTCTGGTGGCCG  
CTGGCCGGCGCCTCGGTGGCCCTGCTGCCGCAATCCATGGAGAAGTTCCCCTGGGCGATATCG  
30 GCGACGGTGGCGCGGTGCGGGGTGACGGTGATGCATTTTCGTACCATCGATGCTGATGGCCTTC

50

55

5

10

15

20

25

CTTCAGGTGGTGGCGGGCCGGCCCGAGATGGCGGACCAGATGAAGGGCCTGCGCTACGTCTTC  
TGCAGCGGCGAGGCCCTGGCGCCGGCCACGTGTCAGCCTTTCAGGAGCACATCAACCGAGCG  
GGCAGCATCAGCTTGACCAACCTCTATGGACCCACCGAGGCGGCGGTGACGTCAGCTACTTC  
GACTGCCCCCGGCGCGTCACTCGCGGGGTGCCGATCGGACGAGCGATCACCGGCATCCAG  
5 CTGCTGGTCATGCGCGACGGCGTGCTCAGCCGCCCCGGCGTCGAGGGTGAGCTCGCCATCGGC  
GGCGTTGGTTTGGCGCGCGGCTACATCTCACGGCCAGACCTGACCGCCGACCGGTTCTGTGCCG  
15 CATCCAGGCGGCGACGGCCAGCGGCTCTACCGCACCGGCGATCTGGTGCGCAGGGACGCGGAC  
GGCGAGCTGGTCTTCTGCGGCGCATCGACCATCAGGTGAAAATTCGCGGTCTGCGCATCGAG  
CCCCGGGAAATCGAGGCCAGATCAACGCCCATCCCGATGTGGCCGACTGCGCGCTGATTATC  
20 10 GAGCAGGACTCGGAAACCCTGCCCCAGCTGACCGCCTACATTGTCTGGCGCGACCGGGCTTG  
ACCCGGAAGGCGCTGCTACAGTTCTTGGGCGCGCGGCTGCCCGACTACATGCTCCCGAACCGC  
TTCCTGACCCTCACGGAGCTGCCCGTGACCGCCAACGGTAAGCGCGACTGGCGCGCGCTGCTC  
GGCCCGCTCGAGACCCTGCCTCTCCCTTTCTCC

## 15 (2) peptide sequence

Seq ID No 38 (&gt;pEPOcos6\_ORF2.pep)

30

35

40

MTSAVPTRQTSLLDDFERVADVDPERIAVHASETSLRYGDMNARANRIAHGLRARGIGPNQIV  
AVAMARTPELMIVLYGILKAGAAAYMPIARDAPPLRRDHMLRESQAALMIADEEIIAGLAARVLT  
PADPFFAAMPDHNPEPRHDPTDLIYVIYTSGSTGQPKGVAMEHRAVWNRLTWMQAQYPIDTQD  
20 VILQKTPIVFDVSVWELFWWPLAGASVALLPQSMEKFPWAI SATVARCGVTVMHFVPSMLMAF  
35 LQVVAGRPEMADQMKGRLRYVFCSGEALAPAHVSAFQEHINRAGSISLTNLYGPTEAAVDVSYF  
DCPPGASLARVPIGRAITGIQLLVMRDGVPOPPGVEGELAIGGVGLARGYISRFDLTADRFVP  
HPGGDGQRLYRTGDLVRRDADGELVFLGRIDHQVKIRGLRIEPGEIEAQISAHPDVADCALII  
40 EQDSETLPKLTAYIVVARPGLTRKALLQFLGARLPDYMLPNRFLTLTELPVTANGKRDWRALL  
25 GPLETLPLPFS

45

50

55

5

103

**pEPOcos6\_ORF3 sequences:**

10

**(1) nucleotide sequence**

Seq ID No 39 (&gt;pEPOcos6\_ORF3.seq)

15

5 ATGTTACACCCGATTCCCACCGACCGTTTCGCCCTGAGCCGACCGCTCTTTCGCGGGTACCTC  
GCGCACGATCCGATCGTGCAGGGCGTGCTGGCGGGCGACCATCCAGGCTGGGTCTTGGTGGAC  
CGCGAGCCCGAGCCGCGCACGGCGCTGCTGTGGGCCTTTTCCGATCGGCTCTTCTGCGTGGGC  
GCAGCTGACACGCTGACCCCGCACGCGCTGGCCGAGCTGTTCCACGACCGACTGATCCCCAG  
GCCCCGTAAGATCGGGCAGCCGTTTTCAGGTTTCAGGGCGAGACGGTCGACACCTGGTCCGGAC  
20 10 CACCTGCATCAGGTGTCGCCGACGCGACAGTCTCCTTCCGCCAGGCATTCCGCTTCGACCGC  
GACCTCTTCGAGCGGCTGCCAACCAAGCCGAGCTGGCAGAGGCGCGGCTCGTGCCAATCGAC  
GCGCGGCTGCTGGCCGAACAGGCTGATCTGCGCGAGCGGATACTGGCCTCCTGGTCCAGCGAA  
GCTGCCTTCCATGCGCGCGGTTTCGGCTTCTGCTACCGCGTAGGTGACCAGCTGCCGAGCGTG  
25 TGCCTGGCATCGCACGTAGGCGCGCGCGCGGCCGAGCTGAGCATCAACACCGAGCTCGAAGCG  
CGCAATCGAGGTATGGCAACGCGGCTGTGCCGGCGTTTCATCGCCGAATCGCTGCAGCGCGGC  
CTGACGCCTTGCTGGGGCACCGAGACCTTTCGCCTGCCGTCAATCGCGCTGGCCCAGAAGCTC  
30 GGTTCATCCCGACCTTCACCTTCCCCACCTACTGCTTCGCGACCGGCACCGAACAGCCGGAC  
GACAACTTCCTAGGCGAGCTGTACTACAGGGAATCGCGCATCGCCGGAAGTGGGACCGATGAG  
CCGCAAGCGGTTTCGGCTGGCGCGGGGTTGGAGCCTGGCCGGCGACACCGAGCGTGCCGCGAGC  
20 TTCGCCGCACGCGCCCTGGCCGAAGGGTGGGCGGCCACTCGACTCTGGCCACCGATCCGGAT  
35 TTCGCCCCGATTGCGCGCCAGCGCCGCCTGGCCCCGCCTCAATGTCCCT

35

**(2) peptide sequence**

Seq ID No 40 (&gt;pEPOcos6\_ORF3.pep)

40

25 MLHPIPTDRFALSRLFRGYLAHDPIVQGVLAGDHPGWVLVDREPEPRTALLWAFSDRLFVCG  
AADTLTPHALAELFHDRLIPQARKIGQPFQVQGETVDTWSDHLHQVSPHATVSFRQAFRFDR  
DLFERLPTKPELAEARLVPIDARLLAEQADLRERILASWSSEAAFHARGFGFCYRVGDQLPSV  
45 CLASHVGGGAELSINTELEARNRGMATRLCRRFIAESLQRGLTPCWGTETFRLPSIALAQKL  
GFIPTFTFPTYCFATGTEQPDDNFLGELYRESRIAGSGTDEPQAVRLARGWSLAGDTERAAS  
30 FAARALAEGWAGHSTLATDPDFARLRASAAPRLNVP

50

55



5

10

**pEPOcos6\_ORF4 sequences:****(1) nucleotide sequence**

5 Seq ID No 41 (&gt;pEPOcos6\_ORF4.seq)

15

ATGATTTGTCACTCCCACCGCTTCATTTTCCTCCACGTTCCCAAGGTCGCCGGCACAAGCGTC

AAGGACGTCCTCGGCCAAGAGCTATTCCAGGAGGACCAGGTCACGTTCCAGATCGCTCCCAAT

CCCCACTACCCACCTGAATGGACTGCGCCTTACGAGGAGCACATTATTGCCGCTGAATTGAAG

AGCCAGTTGGCGCCGGAATTTGGGACGATTACTTCAAGTTCGCCTTCGTGCGCCATCCGCTC

20

10 GACTGGGCGGTCTCCAATTACTTCTTCTTCTGCGCGACCGCAAAGGCCATCCGGCCCCACGAA

TTCCTGGAGCGGAAGGGCTTCGCCGGTACCATGGACATGTTTTTCGGAGCGGCCGGCGCCAT

CCGCTGGTCGCCGGCATGCGCTTCAGCCAATGGGAGTTCTTGTGCGACAGCGAGGGCCGGACG

CTGGTGGACTTCGTTGGCAAGTACGAGCGGCTCGAGCAGGACTTCGCCGCCGTGTGTATCCGC

25

ATCGGGCTGACCCCGCCCGACTTGCCGTGCCTCAACCAGACTCGCCACCAATCCTTTACCAGT

15 TACTACGACGAGGCTTTGATGCGCCAAGTCAGCCGCGCGTTAGCTCGCGATTTTCGAAATTTT

GATTATGCC

30

**(2) peptide sequence**

Seq ID No 42 (&gt;pEPOcos6\_ORF4.pep)

20 MICHSHRFIFLHVPKVAGTSVKDVLGQELFQEDQVTFQIAPNPHYPPPEWTAPYEEHIIAAELK

35

SQLAPEIWDDYFKFAFVRHPLDWAWSNYFFFLRDRKGHPAHEFLERKGFAGTMDMFFGAAGRH

PLVAGMRFSQWEFLCDSEGRTLVDFVGKYERLEQDFAAVCIRIGLTPPDLPCLNQTRHQSFTS

YYDEALMRQVSRALARDFEIFDYA

40

**25 pEPOcos6\_ORF5 sequences:****(1) nucleotide sequence**

45 Seq ID No 43 (&gt;pEPOcos6\_ORF5.seq)

ATGAAAGTGGACAAGCGGAATGTCGACGACATTCTCGGACTCACTCCGACACAGACAGGCATC

30 TTGTACCACTACCTGCTGGACCCGCAGGCCGACGCCTATTTTGAACAATTGACGCTGCACCTG

50

55

5

105

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GAGGGGCCGCTCGACGTAGCGCGCTTCCGCCGCGCCTGGGAGCGCGTGGTGGCGGCTCACGAC  
CAGCTGCGCGCCGTGTTTCGCTGGCAAGGGATCGAACACCCGGTGCAGATCATCCTCAAGCAG  
CACGTGCCGGACCTGGAGTTGGCGGAGGTCCCGCGCGACGCCGATCCGGCAGCCTTCCTGGCG  
CAATGGGTGCGGGCCGACCGGGCGCGCAAGTTCGACTTCGAGACGGTGGCCTTTTCGCATCGGC  
CTCTGCCGGA CTGATACCCAACATCACGTGATGCTGCTCAGCAATCACCATATCCTGATGGAC  
GGTTGGAGTACGGGCCTGATTCTGCGGGACTTCCTCGCCTGCTACGGCGACTCCGAAA ACTGG  
CGGCCACGCACCCGAACGCACTTCAAGGCGTTCATCAAGTGGCACCAGAACC GGCCACGCCGG  
GGCGAGGAGCGATTTTGGCGCGACCTGTTGCGCGATGCGCCCGACGGCGGCTTTCCCCGCCTG  
GGCGTCGAAGAAGGCACCCGCCACTCGCTTGACTTCGGCGCCCGCAGCCGCGCTCTCGACGAC  
CGCTTGACCCAAGGCTTGCGCGACATGGCTCGCGACCTCGACGTCACCCTCGCCGCGATGCTC  
CATACCGCTTGGGGCCTTCTACTCCAGCGCTACCAGAACAGCTGCGAAGTGATATTCGGGACC  
ACCGTTTCCGGCCGCAACGTCGAGCTCGCCGGCCTCGACGAGGTGGTGGCTTGTTTCATCAAC  
ACGATTCCGTTCCGCTTCTCGGCCGCGGCCGCGACGACGCCCCGTCGAGGCCTTCGCTGCGGTA  
CAGCGCAATCTGCTGGCGAGAAGCGAGTTCGAAGCCACCCCGCTGGTGGACATCAAGGGCTGG  
AGTGGTCTCGGTCCGGGCGCGGA ACTGTTTCGACACCATCCTGGTCATCGAGAACTATCCCTTG  
GACCGCGCTATCTTCGAGAGTGATTCCAGCCTGCGGTTGACCGACCACCAAATCTTCGAGCGC  
ACCAATTACGGGCTGACCCTGACCATCGAGACCTTCAGCCGGTTGCACGTGACGCTAGCCCAT  
CGCCGTGACCTGCTGGGCGACGCGGCCGCTGAGCGAATGCTAGATCATTTACCGGCCTGCTC  
CAAGCCATGCTGCGCTTCCCTCACCAGCCGTTTCGCGCGCCTCGAGATGAAAAGCGAACACGAG  
GCCCCACCGCTCCTGCACCAACTCAACCAAACGCGTCAGCCGCTGCCGTCCCAATCGGCTTTC  
CACCAGTTGTTCTTCGAGCAGGCCCAGGCCGATGGGGCACGACCGGCGCTGTGGTGGCGCGCC  
ACGCGCTGGACCTACGGCCAGCTGCTGGAACGTGCCCTGCGTCTGGCGGGACGGCTGCAGGAA  
GCCGGCTTCGCCCAGGCGATGTCGCCGCCGTCAGCCTCGGCCCGGTTCCGGATCTGATTCCC  
GGTTTGCTGGGCCCCGCTGTTTCGCCGGCGGCGCCTACCTGCCGCTCGATCCCACCCTGCCGGCC  
CAGCGCTCGCGGTTTCATCCTCGACGATGCCGGTTGCCGCTTCCTGATCAGCGACGCGCCACTC  
GCGGGGCCCACGCCGATCCATCCGGACCCTGCCGGCGCCAGCCCCGTTGACGTCATTTTGGCC  
TGTCAGGACGGCGCCGCGCAGCCCGCCTACCTGATCTACACCTCGGGCTCCACCGGCCAGCCC  
AAAGGCGTCTGGGT TAGCCACCGCAACCTGATCAACTTCCTGACGGGCATGAGCGCAATCCTG  
CCGTCGCGGCCGACGACGTGTTCTCTCGCTGACTACCGTGTCGTTTCGACATTTTCGGGCTC  
GAGACGTGGTTCCCGCTCAGCCGCGGCTGCACGATCGTCTTGGGCACGCGCGCCGAGCAGTTG

5

10

15

20

25

30

35

40

45

50

55

GACCCGGCCGCGGCTGCCAAGGCCATCTCCTGCCATGGCGTCACGGTTTACCAGGCGACGCCA  
TCGCGACTCCAACCTCAACTGGAGCACCCACATTTGTCCGCGCCATCGGCTCCCTGACGACC  
CTGCTGGTAGGCGGCGAACCCCTCCCAGCCGAGCTGCTGCGGCGCGTACGCGAAGTGACCGAT  
GCGCGTATCTTCAACCTCTACGGTCCACCGAAACCACCATCTGGTCCACAGCCGGGGAGGTC  
5 ACCGCGGCGGACGTCCCGGATATCGGCCGCGCGATCGCAAATACCGGCGTTTTCTTCTGGCG  
CGAGACGGCTCGATCCAGCCGCGGGCCTGGTGGGCGAGTTGTGCATCGCCGGCGAGGGCGTG  
15 GCGTTGGGCTACCACCGACGGCCGACCTGAACCGAGAACGGTTTCGCGAGATTCCGCCGGGC  
CGCCTGCCCTTTGCCGGCAAGCTCTACCACACCGGCGACCTGGCCCGCTGGACCGAAGACGGA  
CGGCTCCTCTGCCTGGGCCGTCTGGACGACCAGCTCAAAGTGCGCGGCCATCGCGTCGAGCCG  
20 GGCGAGATCGAGGCAGTGATGGCGCGCCACCCGGCGGTACGCGAGGCGGTGGTCTCACCGCG  
CCGCGCAACGGCGAGCCGGTCTTGGTCCGGTTCTGGACTGCGGAAGGTGAGCCGATGCCAGAG  
GAAGCGCTGAGCGCTTACCTGGCCGACCGACTGCCGAGCTACATGGTACCCGAACGGTGATC  
25 CTCATGAAGGCCATGCCGCTAACCGGCAACGGCAAGATCGACCGGCGCGCCCTACCCAATCCC  
TTCGCCTTGACCGAGTCGACCCGGCAGGCGGCGCCGCGCACCTTGGCCCGCACCGCCGGCGAG  
15 CATCGGGTTGCCGAGCTGTGGCAGGCCTTGTTCGACGCGAGGCGATCGGCTTGGACGAACCC  
TTTTTTCAGGCCGGCGGGAACCTCATTCGGCTTGATTCTGGCTTACGCCAAGCTGGAATCCGCC  
TTCGGGAAGTCGTTCCCGATCACCGATTTGTTCCAGCATACCAGTATTTCGAGCCAGGCAGAA  
30 ATGCTGAGCGGCTCGTCCGTGAGGCGCCGCTCGCGGGAGCCGTGCCGCAACCCCCGGCCGCC  
GCCGCCCAAGTTGCCTCCTCGGCAGCTAAATCCCCAGGGGAGCGCGGCGCGGCAGCGACGTCG  
20 AGCGGCCTGACCGCGCAACCGCCCCAACCCCACTTCCGGCCCATCGCCGTTATCGGCCTCGCC  
35 GGCGATTCCCCGCGCGCACCCGACCTCGACGCCTTCCTTGAAGTCTCACGGAGGGTCGCTGC  
GGCATTCTGCTTCTTCAGCCAAGCCGAGCTGCGCGACGAGGGTCTCGACGCGAATCGAATCGCG  
TGTCATAACTATGTCCCGGCCAAAGGTTTCCTCGACCGGGCCGACCACTTTGATGCCGACTTC  
40 TTCGGCATCCCCGCGCGCGACGCAGAAATCACCGATCCGCAAATTCGGCTTCTGCTTGAGTGC  
25 TGCTGGAACGCGCTGGAGCATGCCGGCTACCCGCCCCGGCGGCGGCGAGATCGGGCTCTTCGCC  
GGCTCCTCGGCCAACTATCACTGGCTCGAATACGTGGGCATTTCCGAGGAGAGCAGCAATCGA  
TTCGCCGTCATGATTCAAAACGAAAAGGACTACCTGGCCACGCGGATCGCCTACCAGCTCGAT  
45 TTGAAGGGCATTGCCGTACCGGTGCAAACGGCCTGCTCGTCGTCGCTGACCGCGGTCGAGCTG  
GCCTGCGATGCGTTACACGCCGGCCGCGTGACCATGGCTTTGGCTGGTGGCGTTGGTCTGACC  
30 TATCCGTTGCGCGCCGGATACCTGCACGAGGATGGAATGATCTTCTCCCCGACGGTCGGTG

5

10

15

20

25

30

35

40

45

50

55

CGGGCCTTCGACGCCCAGGCGGCCGGCACGGTCTGCGGCAACGGTCTGGGCATGGTGGTGCTG  
AAACAGCTCGACGCGGCGCTGGCCGACGGCGATGCCATCCACGCTGTGATTAAGGGCATCGCG  
GCCAACACGACGCGCGCGGCCAAGATCGGCTACACGGCGCCCTCGCAGAACGGTCAGGCGCGG  
GTGATCCGCGCCGCCCATAGGCTCGCCCAAGTCGCGCCGGAGACCATCGGCTATGTAGAAGCC  
5 CACGGTTCGGGCACGCCGCTGGGCGATCCGATCGAGGTGGCGGGCCTGACCGAGGCCTTTGAC  
AGCCCGCGTCGCGGCTTCTGCGCCTTGGGTTCGGTCAAGTCAATGTGGGTCAATTTGGATGCG  
15 GCAGCGGGCATCGCGGTTTCATCAAGGCGGTGCTCTCGCTGTCCCATCGGACCCTGTTCCGC  
AGCCTCCACGTCGACACGCCCAACCCGCAGATCCCGTTCGCCGACGGTCCGTTCAGGTCAAC  
ACGGAGACCCGGCCCTGGCCAGCTGCCGACCATCCCCGCCGCGCCGGCGTCAGCTCCTTCGGC  
20 10 ATCGGCGGCACCAACGTGCACGCCGTCTGGAAGAGGCGCCGAGTTGGCCGAGCACGCGGGG  
CGGCGGCGCGAGCGGCAGCTGTTCTGGTCTCGGCGCGGACTGCAGCCGATCTGGAGCGACGC  
ACCGCGGCGCTGGTCCGCCACCTGGCCGCGCATCCGGACCTCGCACCCAGATGACGTTGCCTTT  
ACCTTGACGCGGGCCGCAAACCGATGACCCACCGTCGTTTCTCTGTCGCCGCCGACCTCGCG  
25 GAAGCCGCCGCGCTCTGGCCGAGCCCGATCCAGTCAAATCCGCCGCGGCGCGCGCCGACCGC  
15 TGCCAGGTCTGGATGTTCCCGGTCTCGGCTCTCAATACCCCGGCATGTGTGGCGGCCTCTAT  
CGCACCGAGCCGGCCTTTCGCGAGCAAGTCGACCGCTGTTTCGACCTCCTCGCGCCGCGTTGC  
30 GATTTGAAGCCCTCGCTCTTCCCCGAGCCCGATCAGGCCATCGACGCATCAGCCCTCGCGGCC  
ATCGACACCGCCCAGATCGCCGTCTTCGTCTGCGAATACGCGCTCGCACGGATGCTGGAAGGC  
TGGGGGCTGCGTCCGGATCGGCTGATCGGTTACAGTTTCGGCGAATACGTGGCCGCTGCCTG  
20 GCCGGCGTCTTCTCCCTGCCCGACGCCTTGGCAATCGTCCGCGAGCGTGGCCGGATCCTGGCG  
35 GCGGCCGAGCCGGGCGCGATGGTCAGCGTGCCCTTCCGGCCGAGCGCGTCGCGTCGCTGCTG  
GAGCCGCCGCTTGCCCTGGCCATTGACAACGGCCCCCTCATGCGTGGTGTCCGGGCCGGTCGAA  
CCGGTGCGCACCTTCACCGCTCGCATGAACGGGACCGGGTCTGGGTGACGCCGCTCCAGGCC  
40 GAGCGCCCGATGCATTCGCCGCTGATGGCCGAGGCCGGCGGCTCACTGCGCGCCATGTTGGCC  
25 GGGTTCCGCTGAATGCGCCGCGAATCCCGATCTTAAGCAATGTTACAGGAACCTACCTAACC  
GACGAGCAGGCCCGAGACCCGATTACTGGGCCCGTCACCTGTGCGGCAACGTTTCGCTTCGCC  
GACGGTGTGCGAACCTTGTTGGCCGAGCGCGATCCGGTGTTCCTTGAATTCGGGCCGGGCCGC  
45 GATCTGAGCTCCTTGGTGCGCCACCAGATGCCGGAAGGCGCCGACGAGCCGATCGCACTGATC  
CGTCATCGCGAAGATCCGGTGCGCGACGAAGACCTCCTGCTCGATGGCTTGGGCCGCTGCTTC  
30 CTGCGTGGGGCGACCCTCCACGGGCAGGCCTTGACGCCGCCGAGGCTGCCGCCGCGTGCCG

5

10

15

20

25

30

35

40

45

50

55

CTGCCCCGGTTACCCGTTCCAGGGTCCACGCTGCATGCCGGCCCCGCGCCGGACTGCCCGGCCTG  
GCGCGACCGACCGTGGGAGCGACCACCATCAGCTACCGACCAGCCTGGAAGCGGGCGCCGCGC  
TTGGCGGCTGTGGAATCGCTCGCGCCGCAATCCTGGTTGGTATTCAGCGACGGCAGCGAATTG  
GCGGGCGAGCTGGTGGCCGGCCTGCGCGCTTCCGGTTGCGCGACCACCCTCGTCGAAGGTGGG  
5 CTGGCGTTCGCGCGCTTCGCGGGCGGCTTCCGCGCGAATCCCCGCGAGGAACAAGATCTCGCA  
CAGCTGTTGCGGACCCTGTCGGCCGAAGCGATGCTGCCACCCACATCCTGCACCTGCTCAGC  
15 CTGCCGTGCGCCGAGCGCGACTCGCCGCTGGCGCGCCTGGAGCACCTCACCGAGCTGGGCTTC  
CACCATCTGCTGGCCCTGGCCCGCCAACTGGAGGCGGTGCGCGCCCCGAGGTCCGCCTCGCC  
GTGGTGACAACCGGCCTGGCGGCGATTGGCGGCGAGTCCGAGCTGCGGGCCCGAGGTGCGGCTG  
20 10 TTGCGGGGACCTGTCCGCGTGATTCCCTTTGAATTCCCGAAGTTGCGGCTGCGCCTGATCGAC  
CTCGACTCGGCCGATCCCATCTGGCGTAGCGGTTGTGAGCCGTTGCTGCGCGAAATGGGCGCT  
GCCCCGGGACCTGAAGAAATCGCGCTGCGCGGCACCAGCCGTTGGGAGTTGGGCTACGAGCCG  
GTCGAGGGGGGACCGTGAGCACCATCTCCTCGCGACTGCGCGAGGGCGGCGTCTATCTGATC  
25 ACCGGTGGCCTCGGCGGCCTGGGTCTGGCCTTGGCCCGTCACCTCGCCCGGAAGTACCGCGCC  
15 ACCCTGATCCTCGCTGGCCGGCGAGGCGCGCCGGCGCGCGAGCTCTGGCACCAGGCGCCAGCG  
GAGTTCGTACCGGTGCGAGCTGCGATCGCACAGATGGAGGAGTGTGGCGCCCGCGTGATTCCC  
30 GTCGCGCTCGACGTCACCGACGCCGACCAAGTGAACGCGTTGTTGCCACCATAGAAGCTACG  
GTCGGAAGATTGAAGGCGTTTTCCACATGGCTGGCATCGTTGACGGCGGCATCATTGGAACG  
CGCACGCGCGCTGCCAGCGACGCCGTGCTGGCGCCCAAAACGGTCGGAACCTGGATTCTCGAT  
20 CGGGCTCTCCGCGGCGCCGGTGGCGGCTTCCCTGGTGCTGTACTCCTCGATCAACGCGGTGCTC  
35 GCGCCCTTCGGCCAGGTTGCCTACGCGCCGCGCAACGCCTTCCCTCGACGCCTTCGCCAGCGCC  
CACGAACACGACGAGCGTCTTTTCCGCGTCAGCATCGGTTGGGACACCTGGCGCGAGGCCGGC  
ATGGCCGTGATGCCGCCCCGCGCCCGCGGCGACAGGCCCGCTCGAAGGGCTTAGCGACGAG  
40 CAGGGCTTGCGCCTGCTCGAAAGCGCCTTGGTGGTTGCGAACCGCGACTCCTCGTCTCCATC  
25 AGCGAACTGCGCGCTCGACTAGCCGAGCATCATCGCAACGGCGGCATTCCCCGGTTGCTCGGG  
CCCCGCGCCAACGAGGCGGGTGCGAGCTGATTCCGGCGAGGAGGGCGCCACGCAAGACGCGTCG  
CCGGCCCGTCGCGCCCGTCCCGATCTGGTCGTGGCCTTCGCGCCGGCGGCAACGAGCTGGAG  
45 CGCCGGATCGTGGCCATCATCGCGGCCTACCTGCGGCTCGGTGAGGTGGGCGTCGACGACAAC  
TTCAACGATTTGGGCGCCACCTCGCTCGACCTCATCCAGATCGCCCAACGCCTCGGTGCGGAG  
30 TTGGGCCGCGATGTCCCTGTGCTCTCGCTCTACCAACACCGCACCGTACGCGGGCTGAGCCGC

5

10

TTCTTCGGCGGCGCGCTCCAATCCGCGCGGTCCGGCGTCCCGACGGGCGCTGCCGCACCGGGC  
GCCGCCACGCCGGGGGTTGCCACCCCGCCGCGGCCACAACCGTCGCGCCAGCACCTGGAAAAA  
CGCCGTCAATTGAGGAAAAAAGGGGGGCCTTCCCATCATGAG

## 5 (2) peptide sequence

15

Seq ID No 44 (&gt;pEPOcos6\_ORF5.pep)

20

10 MKVDKRNVDLGLTPTQTGILYHYLLDPQADAYFEQLTLHLEGPLDVARFRRRAWERVVAHD  
QLRAVFRWQGIEHPVQIILKQHVDPDELAEVPRDADPAFLAQWVAADRARKFDFETVPFRIG  
LCRTDTQHHVMLLSNHHILMDGWSTGLILRDFLACYGDSENWRPRTTRTHFKAFIKWHQNRPRR  
15 GEERFWRDLLRDAPDGGFPRLGVEEGTRHSLDFGARSRALDDRLTQGLRDMARDLDVTLAAML  
HTAWGLLLQRYQNSCEVIFGTTVSGRNVELAGLDEVVGLFINTIPFRFSAAAATTPVEAFRAV  
QRNLLARSEFEATPLVDIKGWSGLGPGAELFDITLVIENYPLDRAIFESDSSLRLTDHQIFER  
25 TNYGLTLTIETFSRLHVTLAHRRDLLGDAAAERMLDHFTGLLQAMLRFPHPQPFARLEMKSEHE  
AHRVLHQLNQTRQPLPSQSASFHQLFFEQAQADGARPALWCGATRWTYGQLLERALRLAGRLQE  
15 AGFARGDVAAVSLGPVPDLIPGLLGPLFAGGAYLPLDPTLPAQRSRIFLDDAGCRFLISDAPL  
AGPTPIHPDPAGASPVDVIFACQDGAAQPAYLIYTSGSTGQPKGWWSHRNLINFLTGMSAIL  
30 PVAADDVFLSLTTVSFDIFGLETWFPLSRGCTIVLGTRAEQLDPAAAAKAISCHGVTVYQATP  
SRLQLQLEHPTFVRAIGSLTLLVGGEPLPAELLRRVREVTDAIFNLYGPTETTIIWSTAGEV  
TAADVDPDGRPIANTGVFLLARDGSIQPPGLVGELCIAGEGVALGYHRRPDLNRERFREIPPG  
20 RLFFAGKLYHTGDLARWTEDEGRLLCLGRLLDQLKVRGHRVEPGEIEAVMARHPAVTQAVVTR  
35 PRNGEPVLVGFWTAEGEPMPEEALSAYLADRLPSYMPERCI LMKAMPLTGNGKIDRRALPNP  
FALTESTRQAAPRTLARTAGEHRVAELWQALLRREAIGLDEPFFQAGGNSFGLIRLHAKLESA  
FGKSPITDLFQHTSIRSQAEMLSGSSVEAPLAGAVPQPPAAAAQVASSAAKSPGERGAAATS  
40 SGLTAQPPQPHFRPIAVIGLAGRFPAAPDLDAFLELLTEGRCGIRFFSQAELRDEGLDANRIA  
25 CHNYVPAKGFLDRADHFDADFFGI PPRDAEITDPQIRLLLECCWNALEHAGYPPGGGEIGLFA  
GSSANYHWLEYVGISEESSNRFAVMIQNEKDYLATRIAYQLDLKGI AVTVQTACSSSLTAVEL  
ACDALHAGRVTMALAGGVGLTYPLRAGYLHEDGMI FSPDGRCAFDAQAAGTVCGNGLGMVVL  
45 KQLDAALADGD A I HAVIKGIAANN DGA AKIGYTAPSQNGQARVIRAAHRLAQVAPETIGYVEA  
HGS3TPLGDP I EVAGLTEAFDSPRRGFCALGSVKS NVGHL DAAAGIAGFIKAVLSLSHRTLFA  
30 SLHVDTPNFI PFADGPFQVNTETRPWPAADHPRRAGVSSFGIGGTNVHAVLEEAPQLAEHAG

50

55

5

10

15

20

25

30

35

40

45

50

55

RRRERQLFLVSARTAADLERRTAALVRHLAAHPDLAPDDVAFTLHAGRKPMTHRRFLVAADLA  
EAAARLAEPDPVKSAARADRCQVWMFAGLGSQYPGMCGGLYRTEPAFREQVDRCFDLLAPRC  
DLKPSLFPEPDQAIDASALAAIDTAQIAVFVCEYALARMLEGWGLRPDRLIGYSFGEYVAACL  
AGVFSLPDALAIVRERGRILAAEPGAMVSVPLPAERVASLLEPPLALAIIDNGPSCVVSGPVE  
5 PVRTFTARMKRDRVWVTPLQAERPMHSPLMAEAGGSLRAMLAGFRLNAPRIPILSNVTGTYLT  
DEQARDPDYWARHLCGNVRFADGVRTLLAERDPVFLEFGPGRDLSSLVRHQMPGAEPIALI  
RHREDPVRDEDLLLDGLGRCFLRGATLHGQALYAGRGCRVPLPGYPFQGPRCMPARAGLPGL  
ARPTVGATTISYRPAWKRAPRLAAVESLAPQSWLVFSDGSELAGELVAGLRASGCATTLVEGG  
LAFARFAGGFRANPREEQDLAQLFATLSAEAMLPTHILHLLSLSPERDSPLARLEHLTELGF  
20 HHLALARQLEAVGAPEVRLAVVTTGLAAIGGESELRPEVGLLRGPVRVIPFEFPNLRRLRID  
LDSADPIWRSGCEPLLREMGAAPGPEEIALRGTSRWELGYEPVEGGTVSTISSRLREGGVYLI  
TGGLGGLGLALARHLARKYRATLILAGRRGAPARELWHQAPAEFVPVAAAIAQMEECGARVIP  
VALDVTADQVNALFATIEATVGKIEGVFHMAGIVDGGIIRTRTRAASDAVLAPKTVGTWILD  
25 RALRGAGGRFLVLYSSINAVVAPFGQVAYAAANAFDAFASAHEHDERLFRVSIWDTWREAG  
RALRGAGGRFLVLYSSINAVVAPFGQVAYAAANAFDAFASAHEHDERLFRVSIWDTWREAG  
15 MAVDAARARGDQAPLEGLSDEQGLRLLESALVGCEPRLLVSISELRARLAEHHRNGGIPRLLG  
PRANEAGAADSGEEGATQDASPARRARPDLVAFAPAGNELERRIVAIIGAYLRLGQVGVDN  
FNDLGATSLDLIQIAQRLGRELGRDVPVVSQYHRTVRGLSRFLGGALQSARSGVPTGAAAPG  
30 AATPGVATPPRPQPSRQHLEKRRQLRKKGGPSHHE

20 pEPOcos6\_ORF6 sequences:

(1) nucleotide sequence

Seq ID No 45 (>pEPOcos6\_ORF6.seq)

ATGAGTGAAGTATCCATTCGCCCCGGCTTGGACATCGCGGTCATCGGCATGGCCTGCCGCTTT  
40 CCCGGTGGCCGCAACCTCGCCGAGTATTGGGCCAACCTGATCGAAGGCCTCGAAACGCTCAGC  
25 TTCTTCAGCGAAGAGGAGCTGCGTGAGGCCGGCTGCGATCCGGTCCAACCTGGCCCAGCACAAAC  
TACGTGCGCAACCAAGGGCCTGCTCCCTGACGCAGACCGTTTCGACGCCGATTTTTTTGGTTAT  
45 TCCCCGCGCGAAGCCCAGGTGATGGACCCCCAGATCCGCGTCTTCCACGAGGTCTGTTGGCAG  
GCGCTGGAGCACGCGGGCTACAACCCGCATCGCCACACCGGCACGATCGGCCTGTTGCGCCGC  
30 GCCGCGCCCAACGTTTTTTGGGAGTTTCTCTCTATCGGTCCGATGCCGCCAATTTAGGCAAC



5

10

15

20

25

30

35

40

45

50

55

TTCACGCTGGGCCTGCACAACAACAAGGACTACCTGAGCTCGCGCATCGCCTACAACTTCAAC  
CTGACAGGGCCCAGCTACACCCTGTTACCGCCTGCTCGACCTCGATGGTCGCCATCCACCAG  
GCCGTCCAGGCGCTGCTCAACGGCGAATGCGACCTGTGCATGGCCGGCTCGGTCTCCATTACG  
CTGCCACTGGTTGCCGGCTACACCTACACGCCGGGCATGATCGTCTCGCCCGACGGCCATTGC  
5 CGCACCTTCGACGCGAGGCGCCAATGGCACTGTCTACGGCGACGGGGCCGGCGTGGTCGTTCTC  
AAGCGGGCCGAGGATGCGTTGGCCGACGGCGACCACATATTTGCGCTCATCAAGGGCTCGGCG  
15 CTCAACAACGATGGCAGTCGCAAGACCGGCTACACCGCGCCCAGCGTGCAGGGGCAGGTGGAG  
GTGATCCGCGCGGCGATGAACCTGGCGGAGGTCGAGCCGGAGGCGATCAGCTACGTGGAACC  
CACGGGACGGGCACCACGGTGGGCGATCCGCTGGAGTTCGAGGCGCTAAAGGAGGCCTTCGGA  
20 GGTGGCTGCAAGGCCTTCTGTGGATTGGGTTCGGTCAAGCCGAACATCGGCCATCTGGACGTG  
ACGTGCGGGGATCGCGAGCTTCATCAAGCTGGTCCTGGCGCTGGAGCACCGCATCCTACCGCCC  
ACGCTCCACTTCCAAC TGCCCAACCCGAAGATGGATGTGGTCGATAGCCCCTTCTACATCGTG  
GCTGAGCGCGAACCCTGGCGCGAAGATCTGCTGCCGCGTCGGGCCGGTGTCAGCGCGTTCGGT  
25 CTGGGTGGCACCAACGTCCACATGATTTTGGAGGAGTTTCAGCGCGAACC GGCGGCGAACAGC  
15 GCGCGCACGCGCCACCTGACGGTGCTGACGGCGCGGTGCGCCGAAGCCCTGGCGCAGCTGGCG  
GCCAACCTCGCCGAACACCTGCGCGAACACCCCGAGTTGGCGCTGGCCGATGTGGCCCATACG  
30 CTGCTGCACGGCCGCAAGCCACATCCATTCGCGCGCATCCTGGTGGCGACCGATACGACGGCG  
GCGATCGACGCCTTGATGAACGACCGCGATCCGCGAACCGGTTTCTTGAAGCGACCGGGCGC  
GGCGAGTCGGTGATCCTGTGTTTTGACGAAACGCCGCGGAGCCGCGAAGCGCCCGCTACCTC  
20 TGGGATCACGAGCCGCTTTATCGCGCGGCGGCGACGTCGTGCTTGGCTGGTGAGGTGCGCCGAC  
35 CCGGATCTGGAAGGCTGCTTTACTGCCCTGATCGCCGAGCAGGGCGCGGCAGCCGCCTTTTGC  
CACCAATACGCGCTGGCCGGATGGCTGCTGGCCATGGGGTTGACCCCGTCGGCGTTGATCGGC  
GTGGGCCAGGGCGAGTGGGTAGCAGCGGCGCTCGCGGAGGTGTTCCCGCCATCGGCCTGCTTG  
40 CGCTGGATTAGGTTCGGCGAACGGCTCCCGCAGCCGCGCGATCAACGGATTCCGTTTCTCTCC  
25 AATTTCTCTGGAAACTGGATCGTTGGGCGTGAGTTGGCCGACCCGGATTACCCAGAAAGCAG  
AAGGGTAAGCGCTGCATGAAGCGCCGTCGGTCCCAACCTCGGTCAGCTGGTGCAGGATGGGGG  
CGATGGAACCGGCTCGGTCAGCTCGTCGCGCGCTGCTCTTCCGCGGGAAGCGGAGGCGGGACG  
45 GTGATCGGCCCCGAGGGCGAGGTT CATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCCAG  
TACCTGGGGGCGAGCTCGAGG

30



5

## (2) peptide sequence

10

Seq ID No 46 (&gt;pEP0cos6\_ORF6.pep)

15

20

25

5

10

15

MACRFPGARNLA EYWANLIEGLETL SFFSEEELREAGCDPVQLAQHNYVRTKGLLPDADRFDAD  
DFFGYSPREAVMDPQIRVFHEVCWQALEHAGYNPHRHTGTIGLFAGAAPNVFWEFLSYRSDA  
ANLGNFTLGLHNNKDY LSSRIAYNFNL TGPSYTLFTACSTSMVAIHQAVQALLNGECDLCMAG  
SVSITLPLVAGYTYTPGMIVSPDGHCRTFDAGANGTVYGDGAGVVVLKRAEDALADGDHIFAL  
IKGSALNNDGSRKTGYTAPSVQGGVEVIRAAMNLA EVEPEAISYVETHGTGTTVGDPLEFEAL  
KEAFGGGCKAF CGLGSVKPNIGHLDVTSGIASFIKLVLALEHRILPPTLHFQLPNPKMDVVDS  
PFYIVAEREPWREDLLPRRAGVSAFGLGGTNVHMI LEEFQREPAANSARTRH LTVLTARSPQA  
LAQLAANLAEHLREHP ELALADVAHTLLHGRKPHPFARILVATDTTAAIDALMNDRDPRTTRFF  
EATGRGESVILCFDET PPEPRSARYLWDHEPLYRAAATSCLAGEVADPDLEGCF TALIAEQGA  
AAAFCHQYALAGWLLAMGLTPSALIGVGQGEWVAALAEVFPPSACL RWIRFGERLPQPRDQR  
IPFLSNFSGNWIVGRELADPDYPRKQKGKRCMKRRRSQPRSAGAGWGRWNRLGQLVARCSSAG  
SGGGTVIGPRARFIS SSTSRRARVRAQYLGASSR

## pEP0cos6\_ORF7 sequences:

30

## (1) nucleotide sequence

Seq ID No 47 (&gt;pEP0cos6\_ORF7.seq)

35

40

45

20

25

30

ATGGAACCGGCTCGGTCAGCTCGTCGCGCGCTGCTCTTCCGCGGGAAGCGGAGGCGGGACGGT  
GATCGGCCCCGAGGGCGAGGTT CATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCAGTA  
CCTGGGGGCGAGCTCGAGGTAGCGGTCCCGCGGCCAGTAGGGCATCGCGCGAATGACGTCGGC  
CAGGTAGGCCTCCGGGTCGAGCCCGTG CAGCTTG CAGCTCGCCACGAGCGAGAAGAGGTTGGC  
CGCGGCGGAGGCGTGTCGTCGCTGCCGAAGAAGAGCCAGGACTTTCTCGCAACCGCAATGGA  
TCGCAGCGCTCGCTCGCTGGCGTTGTTCTCCAGGCGCAGCCGACCGTCGTCGAGGAAGCGCCG  
CAACGGCTGCTCTTGTTGAGGGCGTAGCCGAGCGCGGTGGAGACCAGGCCGCGCTCGCGGGG  
ACGAGCGTGCTCGGCCCTGGCCCAGGCAAAGAACGCGTCGACCAGAGGGCGGACGACGACATC  
GCGACGCACCTTGCGCTGCGCGGGCGGCAGGTCCGCCAGCGCGCGATCGGCGGCAAAGAGGGC  
GTTGATGCGCCG CAGCCCCTCGACACCGAGCTCGTGCTTG CAGACCGCCGCCTCCCAGAAGTT  
GGTACGGCAATGCGACCAGCATCCGACTTCGGTCGGGGGCGGACCGCGCTTCTCGTCGGCAGC

50

55

5

10

15

20

25

30

35

40

45

50

55

AGCGCCTCTTGGTGGTGTGCCGCGGAAGAGGGCGTCATAGATGGCGTGAGCGTCAGCTTGAAT  
ATACCGAGAGAAGCCGCGGAACATCTCGCAGACCGCGGCGCTGGTATGCTTGGGCTGGTACTC  
GAAGAAGACGTGATCCTTGTCCGCGAGGACGACGAAGAAGTGTCCCTTGCGGCACGGCCCCGGG  
CTTCTTGTCTTTCGCTCCTGGATGGGCCCAGGCTGGACGGAGACCCCGGTGGCGTCCGTGGA  
5 CAGGCAGAAGGCGGTCTCGAAGGCCTCTTTGCGCGCGGCCCTCGACGATGGCGCCCAGGGTCGC  
ACCGACGTCTTCGGCGTAGCGGCACATCGTGCCGCGATCGAGCGACGCGCCCTGAAGCTCCAG  
CTGCTGCTCCAGTCGATAGAACGGGACGCCGAGCAGGTACTTGCTGGTGAGGATGTGCGCAAT  
CATCGACGGCGCGAGGAACGACCGCCGGAACAACCTCCTTCGGAAGCGGCGTCGTGATGAAGAC  
CGTGACAGGTCTCGCCCTTCGGCGCCGGCGGGCGGCGCGTTCGAGCGAGGGCGCGTTCGAGCGCTGT  
10 GGAGGAAGCGCTCGGCTCGCCGGCCGCTGCCGTGTCTTCGGGCTGACGCTCGCAGCGGGCGT  
CGGCGTCGGGGCTTCTCTCGCGACGACCTGGAGCGGGGCGCTTCCTCCTCGCCCCGAAGTCT  
CGCATCCGTGACGGACCGCTCGGCCTTGACACGACGCGTGCGAGCACGATGCGGCGCATTC  
GCCGCGCTCGTAGCCGAGTCGCGAGGTCTCCTCGACCCCGATGCGCGTCCCGTTCGCATCGAG  
CTCGGGGCAGGAGAGCTCGATGCGGACGACGGGCAGGTCCGACTCCGACAGGTCCGACGGCC  
15 CTTGCCGCCGGACCTTCGTTTCGGCCCCCTTGGGGTCGTGCTGCTGCCGCTCGTCGCCTGTATT  
GCGCTCGGCGGCGTCGAGTGCCTTCGCGAGGCGCTGGACCTCGAGGAACATCGAGTCGAACGC  
CAGCTGCTCCGCGCTCACCTCGGCGCGCTCCGCCTTGGCCACGAACAGTCGACGTTCGAGAAG  
CTGCAGCTGCTCGAGCGCACGGGTGTAGGCGCGCCGAAGCTGCGCGAGCGCATCGCGCGCTCC  
CACGAGCTCGCTCTTTGCCGCGGCGAGCTCCGCTTCGAGCTGCGCGATGCGCTGCTGCTCGGC  
20 CGAGAGCGTCGGCTTGGCGGCGGCGTCTGTCACGACGCGCGCTCTACGTAAGCCGCGCGTACTT  
GTCGAGCGAATTTCGTGCGGCTCAGTGGACGCGGCGCGGTGCGCGCCTTCGCGGTTTGGACGTG  
GGCGCGATCTCGATGCCGTCGAGCAGCGTCTCGAGCGTGGCGTCCACCTCGACGTGCGTG  
GCGCCCTCGGTCCGGGGGTCCGGAAGTGCGAACGCTCCGCGATCAAGGCGTTTTGAAAACAGG  
CAGATTCCACTGCCATCGAAGAAGAGAATCTTGATCGTGGTCCGCCGCTTGCCGACGAACGCG  
25 AACAGCGCTCCGCAGCGAGCCTCGTACCCACACGCTCACGGATGAGACCCGAAAGCCGCTCG  
AAGCCG

5

114

## (2) peptide sequence

Seq ID No 48 (&gt;pEP0cos6\_ORF7.pep)

10

MEPARSARRALLFRGKRRRDGDRPEGEVHLVVDEPGAGARPVPGGELEVAVPRPVGHRANDVG

QVGLRVEPVQLAARHEREEVGRGGGVVVAEEEEPGLSRNRNGSQRS LAGVVLQAQPTVVEEAP

5

QRLLLVEGVAERGGDQAALAGTSVLGPGPGKERV DQRADDDIATHLALRGRQVRQRAIGGKEG

15

VDAPQPLDTELVLADRRRLPEVGTAMRPASDFGRGRTALLVGSSASWWCAAEEGVIDGVS VSLN

IPREAAEHLADRGAGMLGLVLEEDVILVREDDEEVSLAARPGLLV LALLDGPRLDGD PGGVRG

QAEGGLEGLFARGLDDGAQGR TDVFGVAAHRAAIERRALKLQ LLLQSIERDAEQVLAGEDVRN

HRRREERPPEQLLRKRRRDEDRAGLALRRRRRRRVERGRVERCGGSARLAGRCRVL RADARSGR

20

10

RRRGFSRDDLERGRFLLARTARIRDGPLGLVHDACEHDAHSAALVAESRGLLDPDARRRRIE

LGAGELDADDGQVGLGQVATALAAGPSFRPLGVVVLPLVACIALGGVECLREALDLEEHRVER

QLLRAHLGALRLGHEQSTSQKLQ LLLERTGVGAPKLRERIARSHELALCRGELRFELRDALLLG

25

RERRLGGGVVHDAALRKPRVLVERIRAAQWTRRGARLRGLDVGAISMPSSSVSSVASSTSTCV

APSVGGSGSANAPRSRRFENRQIPLPSKKRILIVVRRLLPTNANSAPQRASYPTRSRMRPESRS

15

KP

30

## pEP0cos6\_ORF7.1 sequences:

## (1) nucleotide sequence

20 Seq ID No 49 (&gt;pEP0cos6\_ORF7.1.seq)

35

ATGTTCTCTCGAGGTCCAGCGCCTCGCGAAGGCACTCGACGCCGCCGAGCGCAATACAGGCGAC

GAGCGGCAGCACGACGACCCCAAGGGGCCGAAACGAAGGTCCGGCGGCAAGGGCCGTCGCGAC

CTGTCCGAGTCCGACCTGCCCGTCGTCCGCATCGAGCTCTCCTGCCCCGAGCTCGATGCGACG

40

GCGACGCGCATCGGGGTCGAGGAGACCTCGCGACTCGGCTACGAGCGCGGCGGAATGCGCCGC

25

ATCGTGCTCGCACGCGTTCGTGTACAAGGCCGAGCGGTCCGTCACGGATGCGAGCAGTTCGGGC

GAGGAGGAAGCGGCCCCGCTCCAGGTCGTTCGCGAGAGAAGCCCCGACGCCGACGCCCGCTGCG

AGCGTCAGCCCGGAGGACACGGCAGCGGCCGCGAGCCGAGCGCTTCCTCCACAGCGCTCGAC

45

GCGCCCTCGCTCGACGCGCCGCGCCGGCGCCGAAGGGCGAGACCTGCACGGTCTTCATCACG

ACGCCGCTTCGGAAGGAGTTGTTCCGGCGGTTCGTTCTCGCGCCGTCGATGATTGCGCACATC

30

CTCACCAGCAAGTACCTGCTCGGCGTCCCGTTCTATCGACTGGAGCAGCAGCTGGAGCTTCAG

50

55

5

10

15

20

25

30

GGCGCGTCGCTCGATCGCGGCACGATGTGCCGCTACGCCGAAGACGTGGTGCGACCCTGGGC  
GCCATCGTCGAGGCCGCGCGCAAAGAGGCCTTCGAGACCGCCTTCTGCCTGTCCACGGACGCC  
ACCGGGGTCTCCGTCCAGCCTGGGCCCATCCAGGAGCGCAAGGACAAGAAGCCCGGGCCGTGC  
CGCAAGGGACACTTCTTCGTCTGCTCCTCGCGGACAAGGATCACGTCTTCTTCGAGTACCAGCCC  
5 AAGCATAACCAGCGCCGCGGTCTGCGAGATGTTCCGCGGCTTCTCTCGGTATATTCAAGCTGAC  
GCTCACGCCATCTATGACGCCCTCTTCCGCGGCACACCACCAAGAGGCGCTGCTGCCGACGAG  
15 AAGCGCGGTCCGCCCCCGACCGAAGTCGGATGCTGGTTCGATTGCCGTACCAACTTCTGGGAG  
GCGGCGGTCTGCAAGCACGAGCTCGGTGTCGAGGGGCTGCGGCGCATCAACGCCCTCTTTGCC  
GCCGATCGCGCGCTGGCGGACCTGCCGCCCCGCGCAGCGCAAGGTGCGTCGCGATGTCGTCGTC  
20 CGCCCTCTGGTCGACGCGTTCTTTGCCTGGGCCAGGGCCGAGCACGCTCGTCCCCGCGAGCGC  
GGCCTGGTCTCCACCGCGCTCGGCTACGCCCTCAACCAAGAGCAGCCGTTGCGGCGCTTCCTC  
GACGACGGTCGGCTGCGCCTGGAGAACAACGCCAGCGAGCGAGCGCTGCGATCCATTGCGGTT  
GCGAGAAAGTCCTGGCTCTTCTTCGGCAGCGACGACCACGCCTCCGCCGCGGCCAACCTCTTC  
25 TCGCTCGTGCGGAGCTGCAAGCTGCACGGGCTCGACCCGGAGGCCTACCTGGCCGACGTCATT  
CGCGCGATGCCCTACTGGCCGCGGGACCGCTACCTCGAGCTCGCCCCAGGTACTGGGCGCGC  
15 ACCCGCGCCCGGCTCGTCGACGACGAGATGAACCTCGCCCTCGGGCCGATCACCGTCCCGCCT  
ACCGCTTCCCGCGGAAGAGCAGCGCGCGACGAGC

(2) peptide sequence

35

40

45

50

55

20 Seq ID No 50 (>pEP0cos6\_ORF7.1.pep)  
MFLEVQRLAKALDAAERNTGDERQHDDPKGPKRRSGGKGRDLSESDLPVVRIELSCPELDAT  
ATRIGVEETSRLGYERGMRRIVLARVVYKAERSVTDASSSGEEEAAPLQVVAREAPTPTPAA  
SVSPEDTAAAGEPSASSTALDAPSLDAPPPAPKGETCTVFITTPLPKELFRRSFLAPSMIAHI  
LTSKYLLGVPFYRLEQQLELQGASLDRGTMCRYAEDVGATLGAIVEARKEAFETAFLSTDA  
25 TGVSVQPGPIQERKDKKPGPCRKGHFFVVLADKDHVFFEYQPKHTSAAVCEMFRGFSRYIQAD  
AHAIYDALFRGTPPRGAAADEKRGPPPTTEVGCWSHCRTNFWAAVCKHELGVGLRRINALFA  
ADRALADLPPAQRKVRRDVVVRPLVDAFFAWARAEHARPRERGLVSTALGYALNQEQLRRFL  
45 DDGRLRLLENNASERALRSIAVARKSWLFFGSDDHASAAANLFSLVASCKLHGLDPEAYLADVI  
RAMPYWPRDRYLELAPRYWARTRARLVDDMNALALGPITVPPPLPAEEQRATS

30

5

116

**pEPOcos6\_ORF7.2 sequences:**

10

**(1) nucleotide sequence**

Seq ID No 51 (&gt;pEPOcos6\_ORF7.2.seq)

15

5 ATGATTCCGGCGGGCGTGCAGGTGTTTCGTCGCGCTGGAGCCGGTGGACATGCGCTACGGCTTC  
GAGCGGCTTTCGGGTCTCATCCGTGAGCGTGTGGGGTACGAGGCTCGCTGCGGAGCGCTGTTTC  
GCGTTCGTCGGCAAGCGGCGGACCACGATCAAGATTCTCTTCTTCGATGGCAGTGAATCTGC  
CTGTTTTCAAACGCCTTGATCGCGGAGCGTTTCGCACTTCCCGACCCCCCGACCGAGGGCGCC  
ACGCACGTCGAGGTGGACGACGCCACGCTCGAGACGCTGCTCGACGGCATCGAGATCGCGCCC  
20 10 ACGTCCAAACCGCGAAGGCGCGCACCGCGCCGCGTCCAC

**(2) peptide sequence**

25

Seq ID No 52 (&gt;pEPOcos6\_ORF7.2.pep)

MIPAGVQVFVALEPVDTRYGFERLSGLIRERVGYEARCGALFAFVGKRRTTIKILFFDGSIG  
15 LFSKRLDRGAFALPDPTTEGATHVEVDDATLETLLDGIEIAPTSKPRRRAPRRVH

30

**pEPOcos6\_ORF7.3 sequences:****(1) nucleotide sequence**

35

20 Seq ID No 53 (&gt;pEPOcos6\_ORF7.3.seq)

ATGACAAGGACGAAGGCGACCGAAGTGATGTGGTCCGAGCGCGTTCGGGCGTGGCGCGAGAGT  
GGTGAAACGGCGGAGGAGTTCGCTCGGAGCCGCGGATTTGCGGCCTCGACGCTGCACGGCTGG  
TCGAGCCGGCTGTCGCGGGCCGAGCCACCGCGCTTTCTGCGCCTGGTGCCGAAGGCGCCCGCC  
GTGACGAGCAGCGCTGCGGAGCTCGTCGTCGAGGTCGGCGGCGCGGGGTGCGCGTCGCCGCG  
40 25 GGGTTCGACCCCGCGCTGCTGGCGGAGGTGGTCCGTGCCCTCGGCGGAGCGGGGCGA

**(2) peptide sequence**

45

Seq ID No 54 (&gt;pEPOcos6\_ORF7.3.pep)

MTRTKATEVMWSERVRAWRESGETAEFARSRGFAASTLHGWSSRLSRAEPPRFLRLVPKAPA  
30 VTSSAAELVVEVGARVRVAAGFDPALLAEVVRLGGAGR

50

55

5

10

**pEPOcos6\_ORF8 sequences:****(1) nucleotide sequence**

5 Seq ID No 55 (&gt;pEPOcos6\_ORF8.seq)

15

ACTGGACAGCGCAGCCGGGGTGAGACGGCGCTTCGCGCAGCGCTTACGCAGAAGGCGCGCCGC  
GCGCCATTGTTCGGATGCGGTGCGCGACTTCGCCGCCGATCGGCTGTTGCTGGAAGTGGGACAA  
CCACTGGACGTAACGGCTGAAGCGAGCCAACGGCTCCAGCTCGCGCGGGGCGACCTGTTTCGGC  
GCCTACCAAGCGTTGGCCCAGCTCTGGATCTGCGGCGCCCTGGCCGAACCGCCGCGACTGTAT  
20 10 CCCGACGAACACCGCCGGCGCGTGCCGCTGCCGAGCTACCCCTTCGAGGGAAAGCGGTTCTGG  
ATCGAGGGCTCGCCGTTTCGAAACCGCGCCCCGCCCGCGCCTCACCCCAACCGCCGATTTCG  
GGGACATTCTCAAGGGCGACCCGGCGGACTGGTACTATCGGCCGCGTTTCGAAGCGGCGCCG  
25 CTCTTGCCCAGCCCGTTTCGAGAGCGAACCCGGCGATTGGCTGGTGTTCGAAGATGAGCTGGGG  
CTCGGCGCCTGGCTGAGCGAGACCTTGCGCGACAAGGGCGCGCGGCTCGCGACAGTCGTTTCGA  
15 GGCACCGAGTTCGACGCCTGGCGTCACAGCGCTTCAGCTTCGTCCCGATCGACGGGACGAT  
TACCGGACCCTGCTGCACGAGTTGAAGGCGCAGGGCATCGCGCCGGTCCACCTGTGCCACCTA  
30 TGGAGCGTGACCGCCGCACCGGATGCCGAGCAGTTGCTCGACGTCAGCTTTCACAGCCTGGTC  
CATTTGGCGGCCGCTTTGGGTTTCGGTTGGCTACTTCCACGCCATG

25

30

20 **(2) peptide sequence**

35

Seq ID No 56 (&gt;pEPOcos6\_ORF8.pep)

TGQRSRGETALRAALTQKARRAPLSDAVRDFAADRLLLELGQPLDVTAEASQRLQLARGDLFG  
AYQALAQLWICGALAEPRLYPDEHRRRVPLPSYPFEGKRFWIEGSPFETAPAAGASQPADS  
GDILKGD PADWYRPRFEAAPLLPSPFESEPGDWLVFEDELGLGAWLSETLRDKGARVATVVR  
40 25 GTEFRRLASQRFQLRPDRRDDYRTLHELKAQGIAPVHLCHLWSVTAAPDAEQLLDVSFHSLV  
HLAAALGSVGYFHAM

45

50

55

5

118

## pEP0cos6\_ORF9 sequences:

10

## (1) nucleotide sequence

Seq ID No 57 (&gt;pEP0cos6\_ORF9.seq)

5 ATGAAGTTGAACGTGGTCGCCAACCGGCTATTCGACCCCGAGTCGCCCCGAGCGCACCGAGCCC  
15 GCCAAGAGTCTGTTGCTCGCGGTGACCAAAGTCCTGCCGCAAGAGGTGCCAACGTTCGAACC  
CGCGCCATCAGCGTGGACCTGGATCGCTCGTTTCGACGCGGCGGCGCCCGCCTGGGCGGCCAGT  
TTGTTGGTTGAATGCGGCGCGCCCGTCGAGGAAACGGTGGTGACCTACCATGGCGCAGCCCGA  
TGGCTGCGCCGCTTCGATCGCGTTGCGGTGAATGGTCTCGGCCCCGTCCACCCCGATCAACCT  
20 10 GCGCCGCTGCTGCGCGAGCGCGGCGTGTACCTGATCACCGGCGGCCTGGGCGGCGTGGCTGGC  
CAGTTGGCGCGCTACCTGGCGCGGGCCTGCCGGGCGCGGTTGGTGCTCACCGCGCGCCGGCCC  
CTGCCCCGAGCGCGACCAAGTGGGATCGGGAGTCGGCCGTGCTGTCATGGGACGACAAGACGCGC  
25 CAGCGCATCGAGCTGGTGCGCGAGCTGGAGCGGCTGGGGGCCGAAGTATTGGTGGTGGCTGCC  
GATGTCGCCGACGAAGCGGCCATGGCGCAGGCGATCGAGGCCTCACTGGCGCGATTTCGACGCT  
15 TTGGACGGCTTGATCCACGGCGCCGGGATCGTGCGGGTCGCGTCGGGCGGCACGCCGATCGGG  
AGTATGACGCGGGCCATGTGCGAGGAGCAGCTCCGCCCCAAGATGTTGGGCCTCGACGTCGTC  
30 GACCGCCTCCTGCGCGATCGCCGGTTGGACTTCCGCATTGCCATCTCGTCGCTCGCCCCGATT  
CTCGGCGGCCTCGGCCACGTGCGCTACGCCGCCGCCAACCTCTACATGGACGCGTTCGCGACG  
CGCGCCGCCCGCCGGCAACGCGCCTTGGATCGCGCTGAACCTGGCCGAGTGGGAATACGAGGGC  
20 CCGGCTACCTACGACGAGCGGGTGGGCCGTTTCGCTCAAGCAGCTCGAGCTCACCAACGAGGAG  
35 GGTATCCGCGTCTTCCAGACGGTGTGCGCCTTGGCCGCGCGCGGCCCGCTACAGCAGATCATT  
ATTTCCACCGGCGACCTCCAGGCCCCGCTCGACAAATGGATTACATCAAATCCCTGCATCGC  
CGACCGGGGCGGTCCAGCTCAGTCGCCGGACCGCGGCACCCCAGGGCGGTTTCGGCTCGGAG  
40 CGCGCCGCCTTCGAGGCCCGCCTTCGCTGACGCCTGGTGCGACTTCTTCGGGGTTGAAGAGGTC  
25 GACCCGAACAAAACTTCTTCGATCTGGGCGCCAGCTCGCTCGACTTCATCCACCTCGTCAGT  
CGCTTCAGCAAGGCCATCGAACAGCATGTACCGCTCGAGGCCCTGCTCGAACACTCCACCCTG  
CACGACCTCGCCGCCACCTCGCGGGCGACGCGAACACCGACGCCAGCGACGAAGCGCGCATT  
45 CGCCAACGGCTGCAAGGCGCCAAGTCCGGCGACATCGCCATCATCGGCATGGCCGGCCGCTTC  
CCGCTCGCGCCCCGACCTGGACACCTATTGGC3CAACCTGGTCGGAGGCATCGACGCGGTCAGC  
30 TTCTTCAGCGCCGAGGAGTTGCGTGCTGCTGGCGTCACCGCGGCCGAGATCCACCACACCAAC

50

55

5

10

15

20

25

30

35

40

45

50

55

TACGTGCCGGCCAAGGGGCGCTGCCCGACCAGGACTTGTTTCGATGCGGCCTTCTTTCGAATAC  
ACTGCCAGCGACGCCGAGCTGATGACCCGCAAAATCGCGTGTTACACGAGGTCGTGTGGCAC  
GCGCTGGAAGACGCCTGTTTCGACTTCAACGGCGATCACGGCCAGGTCGGCCTGTTTCGCGGGC  
GCCTCGCCGAACCTGTGGTGGCACTTCGTGGCCAGCTTTTCCGAGGCCGCCAAGACGCAGGGC  
5 ATGTTACACCACCACCCTGCTCAACGACAAGGACTCGATCGCGACCCAGATTTTCATACAAGCTC  
GGTCTAAAGGGCCCCGCGGTACCTTGTTACCGGCTGTTCCACCTCGCTGGTAGCCGTTGAC  
GCCGCTGCCGCTCGATCTGGTCCGGTCAATCGGACATGGCCGTGGCCGGCGCGGTCTCGCTG  
ACTCTCCCCGATAAGGCCGGCTACATCTACGAAAAGGGCATGCTCTTCTCGGCCGACGGCCAT  
TGCCGGGCTTTTCGACGCCAACGCCACCGGCATGGTCTTCGGCGACGGCGCCGGCGCGATCGTG  
10 CTCAAGCCGTTGGACGCGGCCCTGCGCGACGGCGACCCGATCCATGCGGTGATCAAGGGCTGC  
GCCACCAACAACGACGGCGACCCCAAAGCCGGCTACACGAGCGTCAGCGCCCAAGGCCAGGCC  
GAGGTGATCCGCTCGGCCAGATCCTGGCCGACGTGGCGCCCGAATCCATCAGCTACGTGGAA  
GCCACGGTACCGGCACCAAGTTGGGCGACTCGATCGAGATCAAGGCGTTGAAGCAAGCCTTC  
25 GCCAGCGACAAGAACGGATTTTTCGGGCATCGGGTCGGTCAAGACCAACCTCGGTACCTGATG  
GCGGCGGCGGGGATGGCCGGCCTGATCAAGACGGTTCTGGCGATGAAGCACCGCCAATTGCCG  
CCATCGCTGCACTGCGACGAAGTGAACCCCGACCTGGAGTTGGAGCGCAGTCCGTTCTACATC  
AACACCCGCCTGCGCGACTGGGTTGCACCGGGCGGGCCGCTGCGGGCCGGCGTGAGTTCGTTTC  
30 GGGATCGGCGGAACCAACGCTCACGTCATCCTGGAGGAGCCGCCGACGCGCGAGAGCGGCACG  
CGCATGCGCCACTGGAAATTATTGATGCTGTTCGGCGGCCAGCGAGGCGGCGCTCGACCGCCAG  
GCCGATAACCTGGCCGACTACCTGGAGCGCCATCCCGAGGCCACCTCAGCGACGTGGCCTAT  
20 TCCCTCCAGACCGGCCGGCGCGTTCCTGGCCTGGCGGCGCACGGTCCTATGCGAGTACCGCGAG  
GACGCGGTGACCAGTCTGCGCGAGCGACAGGCCAAGCGCGTCCAGACAAGTCGCGTCCGCTGG  
GACCACAAGGACGTGGTCTTCATGTTTCCCGGTACGGGCGCCCAGTACCTCAACATGGGCCGC  
GACTTATACGTCATGGAGCCGGTCTTCCGCGAGGTCATGGACCGCTGCTTCGAGTTGCTGGCC  
40 CCTTTGTGGTCCGAGCATCCGCGCCAGATCCTTTATCCGGAGGGCGGGGTGTGACCCCTGCTC  
CACCGGACTGATTACACCCAGCCGATCGTGTTCTGCTTCGAGTACGCCCTCGCCCATTTGCTG  
CTCTCCTGGGGATTGAAGCCGGCCGCGACCATCGGCTACAGCTTCGGCGAGTACGTTTCTGCC  
TGCCCTCGCCGGCGTCTTCTCCCTGGAAGATGCGATCCGTCTGGTGACCGAGCGCGGTCCGCTG  
45 ATGGCGGCTTTGCCCCGGGCGCCATGCTCAGCGTCCCGGTTCCCGAATGCGAGCTGCTGCGG  
CTGCTGGACGGCTTCCACGCCCAATCGGCGGCCCATCTGGCGCTGGCCGTCGACAATGGCGCC



5

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

TCCTGCATTGTGGCCGGCGAGCAGGCCGCCATCTCGGCCTTCGAATCGATGCTTCGCAAGAAG  
CGTCTGTTGACCATGCGGGTCGCGGTTCAGCCACGCCGCTCATTCGCAGGTCATGACCGGCGCG  
ACCGACGCCCTGCGCAGCATCCTGCGGAAGATCCCCCTCTCCGCGCCGACAATTCCCTTCATT  
TCCTGCGTCACCGGCACCTGGATCACTGCACAGCAGGCTACGGATCGCGAGTATTGGGTGAAC  
CACATGTGCGGGACGGTGCGGTTTCGCGGCGGGTCTGACCGAGCTGGGTCAAAACCGCGAGGCG  
GTGTTCTTGGAAGTAGGTCCGGGCCGCGACTTGACGTTGCTGGCCACCGCATCCTGGCCGAC  
AGCGCGGCCGTGTTTCGAGCTGGTCAAGGCGCCCGACGGCGGCGACGACGATGGGTTCTCCTG  
CTGGATCGATTGGCCAAGCTCTGGAGGCTGGGGATTTTCGATTGACTGGGCCGGCTTCTACGCG  
GATGAGCGGCGGCGGAACTCTCGCTGCCGGGATATCCGTTTCGAGCGGCGGCGCTTCTGGATC  
GAGGGCAACCCGCTGGAGATCGCCGCCGGCAGGCCCAATGTCCAGGGGCCGCTGGTCAAGGCG  
TCGGACATCGGCGCTTGGTTCTACGTGCCGCAATGGCGGCGGTTCGGTGCTCGCCGAGCCGGGT  
ACAACGGCGGCGGGCGCCCGCTCACGGCGGAGCAGGCACGCGTCGTGACCGAGCTACGGGCG  
GGATGCGCGTCGGCCGGCTTGGGCAGCGGGGCTGCGGACTGAATGGCGGTGCCCCGTCCGAG  
CGTCCGAAAGAAAGTGTAGCGCCAGCCGGGTTCGACCAGCGCAGCGGCGCAGACCGGCGCGGAC  
TGCCCGACACCGACTGGGGAGCCAGCGGCTGTGCCAAAGGACGGGGCCGAGCCGCGGCCGACC  
TGGCTTATTTTCGCCGACGCCGGCGGATTGGCCGAATCTTTTCGCCAAGCGGGTTTCAGGCCCGC  
GGCGAGAAGCTTTACCTGGTGGCTTCCGGCTCGCGCTTCGAGCGCCTGGCCGAGACCCGCTTC  
CGCCTCGATCCCGGGGCCAAGTCCGATCACCGCCTGCTTTTCAAGGCGCTCGACGAGGCCGAC  
ATCCTGCCGACCCACCTCCTCGACTTCCGCTCGCTTGAATGCGGCGGGCCCGACGCCGACCCC  
ATGGACCAGGCCGGCTTCTTCGGGCTGTTGCACCTGGTCCAGGCGATGGCAGAGGCCGGCTAC  
AGCCATCCCATTTCGGCTGCTGATCGTCAGTTGCGGCGTCTACGATGTCACCGGTGCCGAACCG  
CTGCAGCCGGCGCGGGCCACGATGATCGGACCGGCTCTGTGCATCCCGCAACAGTATCCGCAC  
CTCGAAACGAGCCATGTGGATTTGGGCGTGGTCCATGCCGACGAGCTCCACGCCGCGCGCCAG  
CTCGACAGCCTACTTGCCGAATGCCTAAGTGCAACGGCCGAGCGCCAATTGGCGCTGCGCGGC  
CGACACCGCTGGCTGCTGGACTACGAGCCAGTCCGCTTGCCGCCGCTCGACCCGGGCCGTCTG  
CCCTGGCGCCAGCGCGGGGTCTACTTGATCACCGGCGGTTTGGGCGGGATCGGCCGCATCCTG  
GCCGAACACCTGGCCCGCACGACCTCGGCTCGCCTGGTCCCTAATCGGCCGCGAAACCTGCCC  
GACCGCGACGACTGGGACGCCTGGCTGAACCGCCCGCAACCGGTTCGACGCCACCCACGAACGG  
CTGCTGCACAAGATCCGCGCGATTTCGCGATCTGGAAGCGCTAGGCGCCGAAGTCTTGGTCCTC  
GCC3CCGACGTCGCCAACGAAGCCGCCATGCGCGAGGCTACGATCGCGCCGAATCCCACTTC

5

10

15

20

25

30

35

40

45

50

55

GGCACAATCCACGGGGTGATTACGGCGCCGGCCTGATGGACGCGCAAAGCTTCTCACTGATC  
GACGCCCTCGACCACGACCTCTGCGCCCGCCAGTTCTGAAGCAAAAATCCGCGGCGTCTGCGTG  
CTCGACCGCGTTCTGGCCGACCGCACGCTCGACTTCTGTCTGCTGATGTCTTCCATCTCCACC  
GTGCTCGGCGGCCTGGGCTATTTTCGGTTACGCCGCGGCCAACGCCTTCCTCGACGCCTTCGCC  
5 CAGGCGCGCAGCCGCGACGCCGCTTTCCCTGGCTTAGCGTGGCCTGGAGCGATTGGAAGTAC  
TGGACCGAGCGCAAGATGGACAACGAGGTCGGCGCCGTCATCGACAGCCTCTCGATGGAACCC  
GCCGAGGGCTTCGAAGCCGTCACCCGCGTCTTGGCTTGGGGCAAGGCGCCCCACATCGCCAAC  
TCGCCCCGTGACCTCGGTGCGCCGCGGGATCAATGGGTCAAACCTGGCCAGCCTGAAATCGGCG  
CACTCCAGCGAGCCCCGAGCCGGCTAGGCATGGACGTCCGGCGCTCTCCAGCGAATGGGTGCGG  
10 CCGCGCAACGTGGTCTGAAGAGAAGCTGGTCCCATTTTCGAGCAGGTGTTCCGGCACTGCGGCA  
CTGGGCATCGAGGACAACTTCTTTGAGTTGCGCGGCGACTCGCTCAAGGCGGTCATGACCGCG  
GCCCCGTATTCAAAGGAGCTGAACGTGGAAGTGCCGCTGCCGACCTTCTTCCAGATGCCACG  
GTCGCTGGCCTGGCCAGTTCTGTGACGCAAGCCAAGCGCAGCGCGGGAGACGATTTCGGCGC  
25 ACCGCGCCGCGCCACATTACCCGCTCTCGGCTGCCAGGGCCGCCATTACCTGCACTACCGC  
ATGGACCCGCGTTGTACCGCATAACAACGATCCCTTCGCCAACCTGATCGAGGGTCCGCTGGAC  
GTGGATCGCGTGGAGCGCATCCTGCACACCCTCATCCTACGCCACGACTGCTTCCGCACCTCG  
TTCCACTTCCGCGAGGGCGAGCCGGTCCAGGTGATTACGATCGGGTGGACTTCAACCTGGCG  
30 CGGATTACCTGCGCGCCCCGAGGATTTGCCCGAACGGATGCGCGATTTTCATCCGCTCCTTCGAT  
CTGGAGCGACCGCCCCGCCATGCGCGCCGGCCTCTTCGTACCGGGCCCGAGCGCCACGTGCTG  
CTAATCGATTTTCACCACATTATCACCGATGGCGTGTGTTTCGAGAACTTCGTGCGCGAGTTC  
35 GCGGCGCTCTACCGCGGCGAGATCCTGCCCGAGCTGGAACCTCGAGTACAAGGATTTTCGCGGTG  
TGGCAGCATGAGAACCGGGGCCGCGCGCCAACAGCGACCAGGCCGCTACTGGACCGAGCAG  
TTGGCCAATGCGCCCCGGGCGGATCGAGCTAACCAACGATTTCCCCCGTCCAGTCGACGCAGC  
TTCCGCGGCGACCGCGTGCGGACCGTGCTTGATGCGGAGCTCGTTGCTCGACTCAAAGAGCAC  
40 GCGGCGCGCCTCGGCATCACCTCTATAGCCTGCTGCTGGGCGGATTCTCGTTATTGCAGCAC  
25 AAGCTCTCCGACTCGCACGACATCGTCATCGGTTTCGCCCCGTCGCGGGCCGCACCCGGAGCGAA  
CTCCAGGATCTGCTGGGCGCGTTTCGTCAACACCCTGCCGATGCGCCACCGCATCGACCCGACC  
45 CATAACGCACGGGTCTTCTTGGAGCAGGTCCACCAGACAACCTTGGCGGCCCTCAGCTACCAG  
GAGCACCCCTTTTGACGAAATGGTGGCGACGCTCGGGTTCGCCGCGGATCCGGCTCGCAACCCG  
30 ATCTTCGACACGATGTTCTTGCTGCAGAACATGGCCATGGGTGCAACCACCATTTCCCGTCTG

5

10

CGGCTCTCGCCTCACGACACTTTTCACCGCAAGGCATTGTGCGACCTGATGCTACAGGCGACC  
GAGTATGACTGCCACCTGGAGCTGGTGTCTGAGTTCGCCACCGACCTGTTCCGGCTGGAAACC  
GCGCAAGTCTTGCTCGACCGCTACCGCCAAGTCTTGAGTGGCTGTTGGCGTACCCCCATGAA  
TCGATAGACGATTTGACGCTCGCCGGCCACTTTTCGCGAAGTCGAAGTGACGATGTCGGACGAG  
5 GGCGACTTTGATTTCTCAGATTTTGAACCCCGCAACGTGAGAAACCTATGGCGCGCC

15

**(2) peptide sequence**

Seq ID No 58 (&gt;pEPOcos6\_ORF9.pep)

20

10 MKLNVVANRLFDPESPERTEPAKSLLLAVTKVLPQEVPNVRTRAI SVDLDRSFDAAAPAWAAS  
LLVECGAPVEETVVVYHGAARWLRRFDRVAVNGLGPFHPDQPAPLLRERGVYLITGGLGGVAG  
QLARYLARACRARLVLTARRPLPERDQWDRESAVLSWDDKTRQRIELVRELERLGAEVLVVAA  
DVADEAAMAQAI EASLARFDALDGLIHGAGIVRVASGRTPIGSMTRAMCEEQLRPKMLGLDVV  
25 DRLLRDRRLDFRIAISSLA P I L G G L G H V A Y A A A N L Y M D A F A T R A A A G N A P W I A L N L A E W E Y E G  
PATYDERVGRSLKQLELTNEEGIRVFQTVLALAARGPLQQIIISTGDLQARLDKWIHIKSLHR  
15 RPGPVQLSRRTAAPQGGFGSERA AF E A F A D A W C D F F G V E E V D P N K N F F D L G A S S L D F I H L V S  
RFSKAIEQHVPLEALLEHSTLHDLAAHLAGDANTDASDEARIRQLQGA K S G D I A I I G M A G R F  
30 PLAPDLDTYWRNLVGGIDAVSFFSAEELRAAGVTAAEIHHTNYVPAKGRCADQDLFDAAFFEY  
TASDAELMDPQNRVLHEVVWHALEDACFD F N G D H G Q V G L F A G A S P N L W W Q F V A S F S E A A K T Q G  
MFTTTLLNDKDSIATQISYKLGLKGPVTLFTGCSTSLVAVDAACRSIWSGQSDMAVAGAVSL  
20 TLPDKAGYIYEKGM L F S A D G H C R A F D A N A T G M V F G D G A G A I V L K P L D A A L R D G D P I H A V I K G C  
35 ATNNDGDRKAGYTSVSAQQQAEVIRSAQILADVAPESISYVEAHGTGTKLGDSIEIKALKQAF  
ASDKNGFCGIGSVKTNLGHLM A A A G M A G L I K T V L A M K H R Q L P P S L H C D E V N P D L E L E R S P F Y I  
NTRLRDWVAPGGPLRAGVSSFGIGGTNAHVILEEPPTRESGTRMRHWKLLMLS A A S E A A L D R Q  
40 ADNLADYLERHPEAHLSDVAYSLOTGRRVLAWRRTVLCEYREDAVTS LRERQAKRVQTSRVRW  
25 DHKDVVFMFPGGAQYLNMG RDLYVMEPVFREVM D R C F E L L A P L W S E H P R Q I L Y P E G G V S T L L  
HRTDYTQPIVFCFEYALAHLLLSWGLKPAATIGYSFGEYVSACLAGVFSLED A I R L V T E R G R L  
MAALPAGAMLSVPVPECELLRLLDGFHAQSA A H L A L A V D N G A S C I V A G E Q A A I S A F E S M L R K K  
45 RLLTMRVAVSHAHSQVMTGATDALRSILRKIPLSAPTIPFISCVTGTWITAQQATDREYWVN  
HMC GTVRFAAGLTELGQNREAVFLEVGPGRDLTLLAHRILADSAAVFELVKAPDGGDDDGFL L  
30 LDR L A K L W R L G I S I D W A G F Y A D E R R R K L S L P G Y P F E R R R F W I E G N P L E I A A G R P N V Q G P L V K A

50

55

5

10

15

20

25

30

35

40

45

50

55

SDIGAWFYVPQWRRSVLAEPGTTAAGAAVTAEQARVVTEL RAGCASAGLGSGACGLNGGAPSE  
RPKESVAPAGSTSAAAQTGADCPTPTGEPAAVPKDGAEP RPTWLI FADAGGLAESFAKRVQAR  
GEKLYLVASGSRFERLAETRFRLDPGAKSDHRLLFKALDEAD ILPTHLLDFRSLDCGGPDADP  
MDQAGFFGLLHLVQAMAEAGYSHPIRLLLIVSCGVYDVTGA EPLQPARATMIGPALCIPQQYPH  
5 LETSHVDLGVVHADELHAARQLDSLAECLSATAERQLALRGRHRWLLDYE PVRLPPLDPGRL  
PWRQRGVYLITGGLGGIGRILA EHLARTTSARLVLIGRETLPDRDDWD AWLNRQPVDATHER  
LLHKIRAIRDLEALGA EVLVLAADVANEAMREAYDRAESHFGTIHGVIHGAGLMDAQSFSLI  
DALDHDLCARQFEAKIRGVCVLD RVLADRTLDFCLLMSSISTVLGGLGYFCYAAANAFLD AFA  
QARSRDAAFPWLSVA WSDWKYWTERKMDNEVGAVIDSLSMEPAEGFEAVTRVLAWGKAPHIAN  
10 SPGDLGRRRDQWVKLASL KSAHSSEPEPARHGRPALSS EWWAPRNVVEEKLVAIFEQVFGTAA  
LGIEDNFFELRGDSLKAVMTAARIQKELNVEVPLPTFFQMPTVAGLAQFVTQAKRSGRETIRR  
TAPRPHYPLSAAQGRHYLHYRMDP RCTAYNDPFANLIEGPLDVDRVERILHTLILRHDCFR TS  
FHFREGE PVQVIHDRVDFNLARITCAPEDLPERMRDFIRSF DLERPPAMRAGLFVTGPERHVL  
25 LIDFHIIITDGVSFENFVGEFAALYRGEILPELELEYKDFAVWQHENRGRRANS DQARYWTEQ  
15 LANAPGPIELTTDFPRPSRRSFRGDRVRTVLDAELVARLKEHAARLGITLYSLLLGGFSLLQH  
KLSDSHDIVIGSPVAGRTRSELQDLLGAFVNTLPMRHRIDPTH TARVFLEQVHQTTLAALSYQ  
EHPFDEM VATLGFAADPARNP I FDTMFLQNMMAGATTI PGLRLSPHDTFHRKALCDLMLQAT  
30 EYDCHLELVLEFATDLFRLETAQVLLDRYRQVLEWLLAYPHESIDDLTLAGHFREVEVTMSDE  
GDFDFSDFEPRNVRLWRA

20

pEPOcos6\_ORF10 sequences:

(1) nucleotide sequence

Seq ID No 59 (>pEPOcos6\_ORF10.seq)

25 ATGGCGCGCCTGAGCCGCACAGATCTCCAACTCGCCATTCACCAGCGCACCGTGGAGCGCGAA  
TATTGGCGCGCTCTGTTCGAGCGCCATCCGCAACGGTCCAGTTTGCCGGGGGTGCTCACC GCC  
CCGATCGGCGACGAGTCGACCCGCGAGACCTTGTCATTCTGCTCCTCGACGAAGATCCCCTTCGG  
45 CTGAGTAATCGTTCGCCGCAACGCCTGCTCACGGTGTTGGCGGCTGGCCTCGCGGCTTTCCTC  
CACCGCTGCGACGGCGCTGAGCGCTTCACCCTGGGGTTGGCCCTACCGCGCCAAGCCGATGAC  
30 CATCACCCGATCCTCAACAGCTTGATCGCGCTGGGGGTCGCGGTCGACTCGAGTACGACCTTC

5

10

15

20

25

30

35

40

45

50

55

CGCGATCTGCTCTATGCGCTTCGATCCGAATACCACGAGGCGATGCGCCACGCCAACTTTCCG  
CTGGCGACCTGGTGGCGCGGCCTACCCGGCGGAACGGCGCCGTTTCGACGTCGCCCTCAGCCTG  
GACCCCTTCACAGACGGCGATTTCGCTGGAAGACCACGCGATCGGCGCGTTGTTCCGGTTCGCA  
TTGGAGGGTGAGCGCCTCACCTGCCGATTGCGATTGACCCCTGCGCGCTATGACCGTCCCGCG  
5 ATCGAAAACCTCGCCGATCGTTTCGCCCCGCTTCCTCACGCGCCTGTGCCGGGACGCCTCCACC  
GTCATCCAGGCGCTGGACCTTTCGCTGCCAAGCGATGAATCGGTGTGGCGCGTCACTGAAGGC  
GTGCGGCGCGGCTATTTCGCAAGACCTGACGCTAGACCGCGCGTTCCGCCGCCAGGCCGCGCAA  
ACGCCCGATCAGCCGGCGATCACGTTGAACGGGGACGTCCAGAGCTACGCCGAGGTCGACCGC  
CGCAGCGACGCGCTGGCCCCGCCACCTCCGTGCCACGGCGTCGGTCCGGAAACGATTGTGGCC  
10 GTCAACGCCCCGGCGCGGGCCTAATCAGCTGACGGCCCTGCTCGCGGTCCATAAGGCCGGCGGC  
GCCTACCTGCCGATCGATGCCGAGGAGCCGGCTGCCCGCCAGCAATTCAAGGTGCGCGACAGC  
GGGGCGCGGTTGGCACTGGAGCCGTGCCCGGACCAGGCGCTGACCGTCACCGACCTGCCGCGG  
CTCTTCCTGGACGATGCCTCGCTCTTCGCTGACGGCGGGCTCGATGTGCCGCGCGGCGCCGAC  
TCGCTCAATCCGGCCTATGTGATGTACACGTCCGGCTCGACCGGACAGCCCAAGGGTGTGGTG  
15 GTTCCCCACCGCGGCGTGGTCAATCGTTTGAATTGGGGGAGTCCCGTTTCCCGCTGGACGAA  
CGCGACCGAATCCTCCAAAAGACGCCGCTGCTGTTTCGACGTGTTCGGTCTACGAGCTGTTCTGG  
GGCGCATGGAGCGGGGCCACCCTGGACATCCTCGAGCCCGGCGCCGAGCGCGACCCCGACGCA  
GTGGCCAGGGCCCTGGCCGAGCGCGCCATTACCGTATGCCATTTTCGTGCCTTCGATGCTGCTC  
GTCTACTTGGAAGTCATGCGGCGGCACCATGCGCCGCCCGTGCCGGACCGGCTCCGTTACGTC  
20 TTCGTGAGTGGCGAGGCCCTCGAACCGGACCACCTCGCCGGGCTCCAGCAGATTGGTTCGGCGC  
CTCGGCCGCACGATTCCCCTCGTTAATCTGTATGGACCAACCGAGGCCTCGATCGAAGTCTCC  
TGCTTCGCCTGTCCCGCCGACCATGTGCCGCGCCGGATCCCCATCGGGCAGCCGATCGACAAC  
GTCGCACTGCACGTTCTCGACCGGCGCGGCCGTGCCAGCCGCCCTATCTTCCTGGCGAGCTG  
TTCCTGGCCGGCGACTGCCTGGCGCGCGGCTACCTCAACCGTCCCGACCTGACCGCGCTCCAC  
25 TTCGTGCCCAATCCCTTCGGCAACGGCGAGCGCATGTACCACAGCGGCGACTTGGCGCTCGTG  
CGCGGCGACGGCCAAGTGGCGTTTCTCGGCCGCCGTGACCACCAAATCAAAATCCGTGGTCAA  
CGGSTCGAACTGGGCGAAATCGAGAGTCATTTGCGCGGGCTCGAAGGCATCGCCGCCGCCGTC  
GTCCAGGCCGAGTCGCAGCACCATGAAACCCTGCTGCACGCCTACGTGCTACCAACGACGCG  
45 GGCCTCAATGCGGCCCGGCTGCGCGCCGCCCTCGCTCAACATCTGCCCGAGTACATGATTCCC  
30 CAGCGCTTCTCGCGGCTGGCCGAGTTGCCGCTGCTGGCGGCAGGCAAGATCGACCGCGCCGCC

125

5

10

CTCGCGCAACGTGCAACGCCGCTCGCCAGCGGCGCGCCCTTCGTGGAACCCAGCGGGCCCCACC  
CAGCAGCGTATCGCAGAACTGTGGCGCCAGGTCTTAGCGGTGCGCGAAGTCGGCGCCGAGGAT  
CCCTTCTTCAGCATCGGCGGCAACTCGCTCAATGTGCTCAAGCTCAGCGCCGCGCTGAGCGAC  
GCCTTCGCGCGTGACATTCCCATGCCGGCCCTGTTCCAATACGACACCATCGCCGCCCAGGCC  
5 TCCTGGCTCGACGGGCAGGTTGACGAACGGGCCCCAATCCGCCGCGCTCGACCGGCAGGCCGCC  
15 GAGGCGGCGCTGACCCTTCAAGAGACCGTGGCCATTTTGTAGGGATTGATGACGAACCA

**(2) peptide sequence**

Seq ID No 60 (&gt;PEPocos6\_ORF10.pep)

20

10 MARLSRTDLQLAIHQRTVEREYWRALFERHPQRSSLPGVLTAPIGDESTRETLSFVLDEDPLR  
LSNRSPQRLLTVLAAGLAAFLHRCDDGAERFTLGLALPRQADDHHPILNSLIALGVAVDSSTTF  
RDLLYALRSEYHEAMRHFPLATWWRGLPGGTAPFDVALSLDPFTDGSLEDHAIGALFRFA  
LEGERLTCRLRFDPAFYDRPAIENLADRFRFLTRLCRDASTVIQALDLSLPSDESVMRVTEG  
25 VRRGYSQDLTLDRAFRRQAAQTPDQPAITLNGDVQSYAEVDRRSDALARHLRRHGVGPETIVA  
15 VNARRGPNQLTALLAVHKAGGAYLPIDAEPAARQQFKVRDSGARLALGPSDQALTVTDLPR  
LFLDDASLFADGGLDVPRGADSLNPAYVMYTSGSTGQPKGVVVPHRGVVNRLNWGQSRFPLDE  
30 RDRILQKTPLLFDVSVYELFWGAWSGATLDILEPGAERDPDAVARALAERAITVCHFVPSMLL  
VYLEVMRRHHAPPVPDRLRYVVFVSGEALEPDHLAGLQQIGRRLGRTIPLVNLYGPTEASIEVS  
CFACPADHVPRRIPIGQPIDNVALHVLDRRGRRQPPYLPGEFLAGDCLARGYLNRPDLTALH  
20 FVPNPFNGGERMYHSGDLALVRGDGQVAFLGRRDHQIKIRGQRVELGEIESHLRGLEGIAAAV  
35 VQAESQHHETLLHAYVVTNDAGLNAARLRAALAQHLPEYMIQRF SRLAELPLLAAGKIDRAA  
LAQRATPLASGAPFVEPSGPTQQRIAELWRQVLAVAEVGAEDPFFSIGGNSLNVKLKLSAALSD  
AFARDIPMPALFQYDTIAAQASWLDGQVDERAQSAALDRQAEEAALTQETVAIFEGFDDEP

40

25

45

50

55

5

## pEPOcos6\_ORF11 sequences:

10

## (1) nucleotide sequence

Seq ID No 61 (&gt;pEPOcos6\_ORF11.seq)

15

5 ATGACGAACCATGACCATCACGAGGAGAGCAGCGGCCTGGAGATCGCCGTCATCAGCATGGCC  
TGCCGATTCCCGGGTGCTATGGCCTGCCGATTCCCGGGTGCTGCCGATTGCGACGCATTCTGG  
GAAAACCTGATCAACGGGACCTCCTCGATCACCCATTTACGCGACGACGAGCTGATCGCGGCC  
GGCGTTGACGCGCGCGACCTGACGCCGAGTACGTGCGCGCGGCCGCCAGATCGATGACGCC  
GAACGGTTCGACGCGGCCTTCTTTGGGTACTCCCAGCGTGAGGCCGAGCTGATGGACCCCCAG  
10 TTCCGCCTGCTCCATGAATGCGCCTGGTCTGTCTGGAACAGGCCCGGCATCGATCCGCGCGTC  
GAAGCCGCGCCGATCGGGCTGTATGCCGGCGCAGCCGACAACACCTACTGGAACGCGCTCTCG  
TCGCTCGACCGGGGCTCGGCCGAATCGGAGCAATTCGCCGCCGAACAACCTTTGCAACCGCGAT  
TTTCTGTGCACGCTGGTCCGCCCGCGCTCAACCTGAAAGGCCCCGCGGTGGTGGTTCAAAGC  
25 GCCTGTTTCGACCTCGCTGTTGGCGGTCCACTCGGCCTGTCGTGCGCTCCTGACCGGCGAATGC  
CGAGTGGCCTTGGCCGGTGGGGTGGCGCTGCGCTTCCCACGCCCCGAGCGGTATCGCTACGAA  
15 CCTGGCATGATCTTCTCGCCCGACGGGGTGTGCCGGCCGTTTCGACGCGGGCGCTAACGGGACC  
GTGCCCCGGCGAAGGCGCGGGGCTGGTAGCGTTGAAGACGCTGAAACGTGCCCTCCAGGACGGC  
30 GACACGATCCACGCCGTGATTTCGCGCGACCGCGGCAAACAACGATGGTGCCCGCAAGACCGGG  
TTCACCGCGCCCAGCGCCACGGCCAAGCCGAAGTCATTTCGCACGGCGCTGCGCCTGGCCCCG  
20 GTGCCGGCCGAATCGATCGACTACGTGAGGCCACGGAACCGGCACGCCGCTAGGCGACCCG  
35 ATCGAGGTAGCCGGCTTGGTGGAGGCCTTCGCCAGCGAGAAGCGCGGCTATTGCCGGCTGGGC  
TCGGTCAAATCCAACCTTGGTCATCTGGACACTGCTGCCGGCATCGCCGGCCTGATCAAGACC  
GTGCTGGCGCTCGAGCACGCGCACATCCCCAAGTCCTGCCACGTGCCACGCCCAACCCCGCG  
40 GCGCGCCTACACAAGACGCCTTTCCGCATTGCCGCCGACGGGATGGCCTGGCCGCGGCGTATG  
25 GCGACGCCGCGGCGGGCGGGTGAAGTTCGTTCCGGCATCGGCGGCACCAACGTCCACGCGATT  
TTGGAGGAGGCGCGCCCCCGCGCGCCGAGCTGGCGGACGGGCGCAGTCAGGTGTTCTCTTC  
TCCGCCAAGGACGAGGCGGCGCTGGACCGTGCCCTTGCCAACTATGGTGCGGCCTTGGAGAAG  
45 CGCGGCGACCTCGCGCGGGCGCGGTGGCCTGGACGCTCCAAAACGGCCGGGCGCATTTCGAA  
TGGCGAGCCAGCGCGGTGGCATCCGACCTCGACGAATTGGCGGGCGCATTGCGCGGCGAGCGG  
30 CCCGGCGCCGTCAAGAAAAACCGAATGGCGCGCGAGGATAAGCCGGTGGCGTTCTTATGTTTCG

50

55



5

10

15

20

25

30

35

40

45

50

55

GGGCAGGGGAGCCAGTACCGTGGCATGGGCCACGACCTGTACCGCGAAGAGCCGCGTTTCCGG  
CACCACCTCGACGCCTGCCTCGCCATCCTCGCCGAACACAAGCCCGAGATCGACTGGCTGGCG  
TTGCTGGGCTACCGCGACGAGGACGAGCCAACCGACCAGATCGGGACGTCCTCGCAGGGCCCCG  
AGCCGGTCAGCCGCATCGAACCCAGCGGAGCTCCTCGACAGCACCGAATTCGCCCAACCTTTG  
CTTTTCTCCATGTCTACGCGCTCGGTGGCTGTGGCTCGACTGGGGCGTGGACCCACGGCG  
ATGATCGGGCACAGCCTGGGCGAGTACAGTGCTGCATGTATTGCAGATTTCTATGCACTCGAT  
CAGGTGCTGCCCTTCATTCTGACCCGCGGTGAGTCATGGCGCAATTGCGGCGCGGCTCGATG  
TTGGCCGTCAGCGGTGACAGCGTTCTGATGCGCGAGCTGATCGCCGATGCGCTCGATTTGGCG  
GCGATCAACGGCGCTGACCAATTTGTCTGGAGCGGGCCGAGCGAGGCTGTCCAAGCCGCGGGG  
GTCCGACTGCGCGGCGCCGGCCTGCGTGCCACCGAGCTGAACACCTCACACGCGTTCCATTCA  
GCCATGATGGATCCCATTTCTGGAGGAGCTAACGGTTGCCGGTTCGCGACTTCAGGTGCGTGTC  
GGGACGATTCCGGTCGTTTCATGCGTTACCGGAACCTGGTTGACGGCGAAGCAGCTGGCCGAT  
CCGCGCTACCACGCGCGTCACGCGCGCGAACC GGTTGCGGCGGGCCTAGCGACGCTG  
ACAGGGGAGGAGCCGCCGCTGATGCTCGAAGTGGGGCCGGGCTCGACCCTGGCGGCTTTGGCC  
CGCGAGCATTCGAATGCCCGCCTCCCGGTGCTCACCAGCCTGCGCCACGCTCGCCAGGCGACG  
CCCGATCGCCAATACCTGCTCGAAACGCTCGGCTGCCTTTGGCGACACGGGGTTTCCGTCGAT  
TGGGGGGCCCATGCCGGACGTTTCGCGACGCTTGTTTCGCTGCCCGGCTATCCCTTTTCCGGC  
GCGGTGCGCCGCTTAGCCGGCGACCCCTCCGCCTGCTGGCCGGAGCCCGCGCCGTCGCCGCC  
CCGTCGGGAACGCGCCAACCTCAGCGCCGACGCGCGCGACCTCCCGAACACTCCGGAGCCGACA  
TCCGGCGCCGTGTGCGCGATCAAAGCGCCAATCGCCGCGCCGATCCCGGCCTCTATCGCCTC  
TCCTGGCGCCAGGCCGGAACGGCGCGCTCGGTCCGCCCGATCTCGGTCCGCCCCGCGACTGG  
ATCGTCTTCGCCTCTGATTCTCACCTGCTCCAGGCGCTCAGGGCCAATCTCGGGACGCGCGCT  
CAGCGGGTGACGCTGGTGACGCCGGGCCAGGAGTACGCAGCCGAGCCGTCCGGGTTTCGGCTG  
CGGCCGGACAGATCGACGATTACCGCGCCCTGTGGGCGGACTTGGCGCAAACCGGTATTGTG  
CCACGATACATCGCGTTCTTCGCCCCGTTTCATGTACCGGGCGCGCATGGCGGGCGATGCCTCG  
ACCCTGGACGAAGTGCGCGAGGGCGGCTTCCTGCCCCCTGACCCGCTTGATCCAGACTCGCCCCG  
CCAGGCGGACCGAGCGGACTTCTAAGCCTCACGATCGTCACCCCGGCCGCCCTGGCGCTGGGC  
GACGAAGCGACGCGCCCGGAATGGGCAATCCTGCACGGGATGGTCGCCGGCTTAAGCCGCGAT  
TATCCCGAATGGCGCTTCGTCTCGATCGACGGCGGCGACCCATCCCCGCATCGGTGCGAAGGT  
CTGGCCCGCTTGATCGCGCTTCATGCGGTGACGAGGCTGGCCCGACCCGCTTGGCGCTGCGC



5

10

15

20

25

30

GGCCTTCACGCTTGGGTTCACAGTGCAGACGTTTCAGCCGGCCACCATCCCTGGGGCGGGT  
ATGTGGCGCGAGGGTGGTGTGTACATGATAACGGGCGGATTTCGGCGGGATCGGTCTGGCGCTG  
GCCCCGCGCCCTGGCTCGAGAAGCTCGCGCCAAGCTGATCCTGGTCGGCCGAAACCTGCCCACC  
GCGCCGATCGATCTCGAGGCTTGGGACGCGCCGCGTTGATTCTCACCGCCGACGTGCGCGAC  
5 GAAGAGGCCATGCGCCGCGTCTTCGATGCCGCGACGCCCGGTTTCGGCGCCATCGACGGCATT  
CTTCACGCGGCCGGTGTCCCCGGTGGCAGCCTGTTCCGCAACCAATCGGACGCGGCCTTCGAA  
15 GACGTGCTGCACGCCAAGGTTTCGCGGTACCCTCGTGCTGCAAGGCCTGAGGGCAATCGATGCG  
CCGCTGTTGCTGATGTCCTCGCTGGACGCTGGCTTCCCGGTCCCGGTCAGACCGCCTATGCC  
GCCGCCAACGCCTTCCTCGACGCCTTCGCCAGTCTGCGCCGGCGAGAGGGAGAGCCGGTGTAC  
20 AGCGTTGGCTGGGACAGTTGGTGGAGGTGGGCATGGCTGCTCGGGTCGCTGCCCCGATCGGCC  
10 GACGAACGCGGCCGCCTGGCGCGGAGGGGATCAGCCCTCCCGAGGGTTGGCAGGCTTTGAGC  
CGGGCGCTCGCCCTCGACCCCCCCCCACCTGATGATCTCGCGCACCGACCTGACCTCGCGCTGG  
CACAGTCGATCCAGCCCTACGCCGGTCGCCTCGAGCGAACCCGAGGTGGCGCTGCCGCGCTGG  
25 ACCGCATCCGCCTGCCAAGCCGTCATCGAGCGTGTTTGGTGGCAGCACTTCGCCACCGCCGCC  
15 GTGCCTCCCGATGGCAACTTTTTTCGAGCTCGGCGCCAGTTCCTTCGACATCGTCCAGCTCAGC  
GCTCGACTTCAACAACAGTTCGGCCGAGATGTCAGCCACACCGTGCTCTACAGTCATCCCACC  
GTCGCCTTGCTGGCCGGCTACTTCGCCAATGACCCGACGCCGTCCGGTGCTGCTGCCGACGAA  
30 CGCGACGAAGCGGTGCGTCGCGGCCGCGACCTCTTGAAGAGCCGCCGGCGAGGAGTA

35

40

45

50

55

20 (2) peptide sequence  
Seq ID No 62 (>pEPOcos6\_ORF11.pep)  
MTNHDHHEESSGLEIAVISMACRFPGAADCDAFWENLINGTSSITHFSDDDELIAAGVDARDLT  
PQYVRAAGQIDDAERFDAAFFGYSQREAEMLDPQFRLLHECAWSCLEQAGIDPRVEAAPIGLY  
AGAADNTYWNALSSLDRGSAESEQFAAEQLCNRDFLCTLVAAALNLKGPVAVVQSACSTSLLA  
40 VHSACRALLTGECDVALAGGVALRFPRPSGYRYEPGMIFSPDGVCPRPFDAGANGTVPGEAGL  
25 VALKTLKRALQDGTIHAVIRATAANNDGARKTGFTAPSAHQAEVIRTALRLARVPAESIDY  
VEAHGTGTPLGDPIEVAGLVFAFASEKRGYCRGSLVKSNLGHLDTAAGIAGLIKTVLALAH  
45 IPKSCHVATPNPAARLHKTPFRIAADGMAWPRRMATPRRAAVSSFGIGGTNVHAILEEAPPRA  
PELADGRSQVFVFSKDEAALDRALANYGAALAKRGDLAAGAVAWTLQNGRAAFWRASAVAS  
30 DLDELALALRGERPGAVKKNRMAREDKPVAFLCSGQGSQYRGMGHDLYREEPRFRHHLDACLA

5

10

15

20

25

10

15

30

ILAEHKPEIDWLALLGYRDEDEPTDQIGTSSQGSPRSAASNPAELLDSTEFQPLLFSMSYAL  
GRLWLDWGVPTAMIGHSLGEYSAACIADFYALDQVLPFILTRGRVMAQLRRGSMLAVSGDSV  
LMRELIADALDLAAINGADQFVWSGPSEAVQAAGVRLRGAGLRATELNTSHAFHSAMMDPILE  
ELTVAGSRLQVGVGTIPVVSCVTGTWLTAKQLADPRYHARHAREPVRFAAGLATLTGEEPFLM  
LEVGPSTLAALAREHSNARLPVVTSLRHARQATPDRQYLLETGLWRHGVSVDWGAHAGRS  
RRLVSLPGYPFSGAVRRLAGDPLRLLAGARAVAAPSGTRQLSADARDLPNTPEPTSGAVSAIK  
APIAAADPGLYRLSWRQAGTAPLGPPDLGPPRDWIVFASDSHLLQALRANLGTRAQRVTLVTP  
GQEYAAEPSGFRLRPDQIDDYRALWADLAQTGIVPRYIAFLAPFMYRARMAGDASTLDEVREG  
GFLPLTRLIQTRPPGGPSGLLSLTIIVTPAALALGDEATRPEWAILHGMVAGLSRDYPEWRFVS  
IDGGDPSPHRCEGLARLIALHAVDEAGPTRLALRGLHAWVPQCEHVQPATIPGAGMWREGGVY  
MITGGFGGIGLALARALAREARAKLILVGRNLPTAPIDLEAWDAPPLILTADVADEEAMRRVF  
DAAHARFGAIDGILHAAGVPGGSLFANQSDAAFEDVLHAKVRGTLVLQGLRAIDAPLLLMSSL  
DAWLPGPGQTAYAAANAFLDAFASLRRREGEPVYSVGWDSWCEVGMAARVAARSADERGRLAR  
EGISPRQGWQALSRLALDPPHLMISRTDLTSRWHSRSSPTPVASSEPEVALPRWTASACQAV  
IERVWCEHFATAAVPPDGNFFELGASSFDIVQLSARLQQQFGRDVSHTVLYSHPTVALLAGYF  
ANDPTPSGAAADERDEAVRRGRDLLKSRRRGV

**pEPOcos6\_ORF12 sequences:**

**(1) nucleotide sequence**

35

Seq ID No 63 (>pEPOcos6\_ORF12.seq)

40

25

45

30

ATGACCGTGGAGCACGAAACCGGATTCGAAATCGCCGTCATCGGGCTGGCTTGCCGCGTTCCC  
GGCGCTGCCGACGTGGCCGCCTTCTGGCGCAACCTGGTCGAGGCCAAGGAGAGCGTGCGCTTC  
TTCGAGGACCACGAGCTGCGGGCCGCCGGCGTGCCCGAGGAGATCTTGCGCCTGCCCAACTAC  
GTGAAGGCCAAGCCACTGCTCGCTGATGGCGAAGCTTTCGACGCGGACTTCTTCGGGTTCCAT  
CCGCGCGAGGCCGCCTACCTGGACCCGCAAGTTCGGCTCCTGCACGAATGTTGTTGGACCGCG  
CTGGAGGATGCCGGCTACGATCCCGCGCAGTACGCCTACCCGATCGGGTTGTTGCGGGGCGTC  
TCCAGCAATCTCTCGTTCCTGTTGACCGCATCGATCCGCGCGACTCCCCCTGCAGAAGCGC  
TATGTGGCCGAGCTGAACGCGGCCTCCTTCGCCACCCAGATCGCCTACCGGCTCGATCTGAAG  
GGGCCGGCCATTTGATTCAAACCGCCTGTTGACGTCACCTGGTGGCGATTACCTGGCGGCG

50

55

5

10

15

20

25

30

35

40

45

50

55

CAAAGCCTGATCGGCGGCGAGTGCCACATGGCCTTGGCCGGCGGAGCGACCTTGGAGGTCCCC  
AAAAAGCCCGGCTATCTCTACCGCGAAGGCTACATCAACTCGCCGGACGGCCACTGCCGGGCC  
TTCGACGCCGACGCGGCCGGCACCATCTTCGGCGACGGCGTCGGCATCGTCCTGCTCAAACGC  
TACCGCGACGCCCTACGCGACGGCCATCACGTGTACGCAGTGATCAAAGGCTCGGCGATCAAC  
5 AGTGACGGCCATCGCAAGGTGTCTTACACGGCGCGGGCAAGAGCGGTCAAGTGGCGGTGATC  
CGCGCTGCGCTGGCGGCGGCCCAGGTAGAGCCGCAAACCATTCGCTTCGTGAGGCCACGGG  
15 ACCGGCACACTCGCCGGCGATCCGATCGAGGTAGAGGCGTTGACGGAGGTCTTTGCCGAAGCG  
GGTCGCGGTACCTGCGCCCTGGGTTCGGTGAAGACCAACATCGGCCACTTGGATGTGGCGGCG  
GGCGTGGCCGGTTTCATCAAGGCGGTCTTGGCGCTCGAGCGGCGCGTCCTCCCGCCCAGCCTT  
20 CACTTCGTCCGGCCCAACCCGGCCATCGATTTCAACGGGCCCTTCTACGTTTGTGCCAAATC  
GAGCGGTTGACGGAGAACGGGCGGTTGCGGGCCGGGGTGAGTTCCTTTGGCATTGGCGGCACC  
AATGCCACGTGATTCTGGAGGAAAGCGCCGGCGCCGGAGGCGAGACTGCCGGCCGGGAGCCCG  
25 CCAGGCGCGAGTCCGTTCTGTTCCTGTTCCCGCTATCGGCCAAGACGCCGGATGCGCTGGCAGGCCGT  
TGCCACGACCTTGCCGACCACCTGCGGGCGCACCCCGAGCTCCTCCTGGCCGATGTGGCCCTC  
15 ACTCTGCAGATGGGGCGGGCGTCCCTTCGCCTACCGCCATGTGGTCCAGGCTGCGACGGCGGAG  
GAGCTGATTGCGCGTCTGGGAGCGTTCGACAGGAGTCCATCCGCAAGAGGCGGAATCGAGTA  
CAATGGGTGTTGGCAGGCGAGGCGATGTGCTTGACGCCGGTTTGCGGCTGTACGCCGATTGG  
CCGGTCTATCGGGAGCGGGTCGACGTCTGTCTGGCGATCGTCGCCAAGCTGCGCCAAATCGAC  
GGCCGGTCATTCTTACATGAGTGGATCGAGCGACCGCGGAGGTTCTTGCCGAATGGTCGACG  
20 GCGCTGGCGTTTATGTTCCACTGCGCGCTGGCGCAAGCCCTGAGCCAGGCCGGCCTGCACCCG  
35 CAGCGCATGTGGAGCCGTGGGCTGGGCGGACAGGTGCGCGTGGTTTTGGCCGAATCCCTGTGCG  
TTGGAACAAGCGCTGGCGCTGGTGTGTGCCAGACACCGGTTCCCGGCGATGCCACACCTCAG  
CGCGAACGCTTGTTTCGGACACTGGAAGGCTGCCGGTTTCGTCCACCACGATTTTGTATTTCG  
GCAGACAGCTCGGGTCGACCCCTGGACCTCGCCGAATTCGCTCATGTGATTTTGGTGCGGT  
40 GGCCAAAGCGCCTCGCCCAATGAGGCGGAGCTGCGCTCATGGAGCGACGCCGCGCCCGAGCTG  
25 GTGACCTTGGCGATCGGCCCATCTTTCTCGAGGCCGCTCCGGGACGGTGGGTCTGGCGATC  
GACCCCAAGCGACCGATGACCTGTCTTCAGCGCACGGTGGCCGCGTTGTGGGAATGGGGATGT  
45 GACGTGCGCTGGGCTGCGTTCACCTCGTCGACCGGGCGTCGGGTTCCTGCCTACCTATCCC  
TTCGTGCGGGTAATTCCCACGATCGGCGACCCCTTCGCGGAGCAGGCGCGGAGGATGACTTG  
30 ATTGCGGCGAGCGCTTCGCGTCCGCCGATCGCCGCCCGAGCCGTGCGCAAACCTCGGCAGCG

5

10

15

20

25

30

35

40

45

50

55

GAACGCCCACGCGCCCAGTCAAGCATCGCCTCGGCAACCACACCGGCTCCGTCTCATACGTCTG  
GCCAGCGTGGCCGTGGCCACCATTCTCGAAACCGTCCGTGCCTATTTGGGGTTCGCCGCCGTG  
CGTTCCACCGACGCCTTCTTCTGAATTGGGCGCGTCCTCGCTGGATTTGGTCAACCTGGGCCAG  
CTCCTTTCCGATCGTCTCGGCCGCGAGGTTCCGACCCTGCTCCTCTACGACCACCCAACACCG  
5 GACCAGTTGGCGCTGGCCCTGACATCCGCGGCGCTCAGCGCAGAGGCGCCGCCCTTAAGGGGC  
GGTCATCGCGCATCGACTTCCGGCACAGCCGCGAGCTCGGCCGCTCCACCGCACCGACGTTT  
15 CCGGGGACGCTCACTCGCAGCCAGCTTCGTTTCGCGAGCAGGACATCGCCATCATCGGGATG  
GCCTTCCGGGGACCGGGCGCCGACGACCTGGACGCGTTCTGGAACAACCTGGTCTGAAGGGGTC  
GAGTCGATCACCTTCTTTCAGCGAGGACGAGCTGCTGGCGGCGGGCGTCCCCCGGAACATCTG  
20 GCCTCGACGCGCTACGTGCGGGCCAAGGGGGAAGTACTGGGATGATGGATTTCTGAACCGGAA  
TTTTTCGGTTATTCGGCGCGCGAGGCGGGTTCATGGACCCGCGAGTTCCGCGTGTTCCACGAA  
TGCTCCTGGCACGCACTGGAGCACGGCGGCTACGATCCGACCCGATGCGCGGCATCGATTGGC  
25 GTCTACGCCGGCGTGACCAACCACCTGCCTTGGCTGATGCGAACTTTGCCGCACCTGACCGAG  
GAGGAGCAATTCGGCGCGCTGCTCCTCACCGACCGCGAGTTTTTCGCACCGCTGCTCTCCTAC  
15 AAGGTCGGCCTGCGCGGACCCGCTATTTTCGCTGCAAACCGCCTGTTTCGACGTCGTTGGTGGCG  
ATCGGCACGGCCTGTGCGGAATTGCGCGCGGGTGCCTGTCAGATGGCCCTAGCGGGCGGCGTG  
30 ACGGCCAGCATCGAGCGCTGCGGCTACTTCCACCAAGAAGGCTACATCCTCTCGCCTGACGGC  
CACACGCGCAGCTTCGACGCGGCGGCCGCGGCACGGTCTTCGGCGACGGAGTCGGCATGGTG  
CTGCTGAAGCCGCTGGCCCAAGCCTTGGCCGACGGCGACACGATCCACGCGGTGATCAAGGGA  
20 ATCGGCATCAACAACGACGGCGCGCGCAAGGTCGGCTTCACCGCACCTAGCCGGGCGGTCAG  
35 ACCGAGGCGATTTCGGGCCGCGCTGCGCGACGCGGGGTGGCGTCGAACCGCGTCAGCTACGTG  
GAGGCGCATGGAACCGCGACCAGAATGGGCGACCCGATCGAGGTCGAGGCCTTGACCCAAGCC  
TTTCGCGCCGAAGCCGACGGTCCGCTTCCGCCCCGGCTCCTGCCTACTCGGCTCGGTGAAGTCC  
40 AACGTGGGCCACCTGAACGCGCGCGCCGGCGTGGCTGGTCTGGTAAAAACCGTGCTGGCGCTC  
25 CAACACCGCCGCCTGCCGACCAGCCTGTTCTACAGTCGCCCAATCCACACATCGACTTTGCG  
GCCAGTCCGTTCCGCGTGAACGGCCAGACTTCGGATTGGGTTCGCGCCAGAGGGGACGCGGTTG  
CTGGCGGGAGTGAGTTCGTTTCGGTATCGGGGGAACCAACGCCACCTGATCGTCGAGGAGGCG  
45 CCGAAAGCGCTACCGACGACAGCGGCACCTCTGTGACGAGCCGAATGACCTCGACGCGGGC  
GACCGCGACGGGCTAGTGCTGCCGATCTCGGCCCGCACGCCGACCGCCCTGGCGCACATCGCG  
30 ACCAACCTCGCCAATCACCTGGAACGACATCCGACCATCGCCCTGGCCGACGTCGCCCTGACC

5

10

15

20

25

30

35

40

45

50

55

CTTCAGCTGGGCGCTCGCCAATGGCCCCATCGCCACAGCCTGATCTGCCGGAATCGAACGGAG  
GCGATCAAGCTGCTGCGCGCCGTCGTCCACTCCGCGGAGGTGCCGCCAGCTCAGGCGCCGGTC  
TCGGATGCGCCGCGCTGTGTTTTCTTTTTCCCGGCCAGGGCGCCCAATACCCGAGCATGGCC  
CGCGACCTGGTTCGAACTGTCCCGACTTCGCCCTGCACCTGGACCCCTGCCTCGACCAGTTG  
5 GCCGAACCTGCTTCCCGAAGATCCGCGTTGCATCCTGTTTCGGCGATGGCCCCGCGGATCGGCTC  
15 GACCAGACGGCCTACACTCAGCCGCTGCTCTTCTCCGTGTCTACGCCTTGGCGCGCTGGTTG  
GGCGATTTTCGGCATTCGCCCCGATGCGATGATCGGCCACAGCCTGGGCGAATACGTGGCGGCC  
TGCTTGCCCGGGCTTTTCTCGCTGAGCGATGCCCTGCTGCTGGTGAGTGAACGCGGCCGCTG  
ATGGGCTCGGCCGCGCGCGGAGCGATGCTGGCCGTCCCCTTGCCCGAATGGGAACCTGGAGGAA  
20 CGCCTGGAGCTTCTGGCCGACGACCGAATCAGCATCGCGGCGGTCAACACCGCCGAGAGCTGC  
10 GTCATCGCGGGACCCAGCGAGGCGATCGAGCGCTGCGCCCAGCGCTGGGCGCGCAAGGCCTG  
ACCTGTACGCCGCTGCGCACGTCCCACGCCTTCCACTCCGCGATGATGGAGCCGATTGTCGAA  
CCCTTCGGCCATGTCTTGGCACGGGTACCTTCGCGCCGCGCGCGCGCTGGATCTCGAAC  
25 CTCGACGGCAAGCCGATCGATTCCGCGGCGGTGATGCAGCCCGACTATTGGGTGCGCCACCTG  
15 CGCCAACCGGTCCGCTTTCACGAGGGACTCAGTCACCTGTTGGCCGAGGACACCCATGCTTGG  
GTCGAAGTGGGTCCCGGCCGAACCTGTCTCCTTCGTCCGCGGCCACCCGGCCTACCGTCAC  
30 CAGCCAATCGTCAACCCCATGCGCCATGCAGTCGAGTCGACGGGCGACGTGCGCCGGTGGCGC  
CAAGCGCTGGGCGAACTATGGCGGGCCGGCATGCCGTCGCCTGGGAGCGGCAGCGGCGCGGC  
CGGCATGCCGGACGACGTGTGCCGCTGCCGGGCTACCCCTTCGAGCGGCGGCCCTTCGCGGCC  
20 CGAAGACCGGTGGAGCTGGCGCAGCCCGCGCCCAAGGCGGAGCTGGTGAAAAACCCCGATCCC  
35 GCGCGGTGGCTGTACCGCCGCGTCTGGCGCCCTGCCCAGGCTGCGGCCGGCGGACTGGCGGTG  
CAGGCGACCGTTCTGGTCTTCGGCGACGGGTCCGAGCTGTGCCGCGCGGCGGTGCTCAGGTG  
CAGCGCCAGGGGCTGAAGTGCGTCTCGATCACCGCGGGCCGCAATTCGCGCGGGAGAGCGAC  
40 ATGCGCTTCACGCTTGACCCCGCTGATCCGCGCCAGCTCGACCAGCTCTTCGCGGCCCTCGAT  
25 GGCTCAGGCTCGCGGCCGCGGTACGTCTTGACCTGCTGACCCTGAACCCGCCCCCGGATGCC  
TCGGCGATCATCGCTCACAGCTACTACAGCCGATGGCCTTGGCTCATGCCTTGGGCGCCAC  
GAGATCGCGCCTGTCTCGATCACCGTCGTACCGCCGGGGTCTGTCGCCGTCGCGGACGAAGCG  
45 ATTGCGAGCCGCTGCAGGCGCTGATCGTGGGCCCCGTGCCTGGTCATCCCGCAGGAGTTTCCC  
GGGCTCAGCGTTTCGGCTGCTGGACGTCAACGTCGACGATCCGGCACCGCGTCTGGCGGAGCGG  
30 CTCGTGGCCGAGCTCTCGGGCACGGATCACATGGTGGCGCTGCGCGGCGGCGAGCGCCTAGTG

5

10

5

15

20

10

25

15

30

35

Seq ID No 64 (>pEPOcos6\_ORF12.pep)

MTVEHETGFEI AVIGLACRVPGAADVAAFWRNLVEAKESVRFFEDHELRAAGVPEEILRLPNY

40

25

45

30

50

55

5

10

15

20

25

30

35

40

45

50

55

ELIRGLGAFRQESIRKRRNRVQWVLAGEAMSLDAGLRLYADWPVYRERVDVCLAIIVAKLRQID  
GRSFLHEWIERPREVPAEWSTALAFMFHCALAQALSQAGLHPQRMWSRGLGGQVGVLAEELS  
LEQALALVLCQTPVPGDATPQERELVRTLEGCRFRPPRFLISADSSGRPLDLAEFAHVDFWCC  
GQSASPNEAELRSWSDAAPELVTLAIGPSFLEAASGTVGLAIDPKRPMTCVQRTVAALWEWGC  
5 DVRWAAFTSSTGRRVPLPTYPFVRVIPTIGDPLRGAGAEDDLIAASASASAGSPPEPSANSAA  
ERPRAQSSIASATTPAPSHTSASVAVATILETVRAYFGFAAVRSTDAFFELGASSLDLVNLGQ  
15 LLSDRLGREVPPTLLLYDHPTPDQLALALTSAAALSAEAPPLRGGHRASTSGTAASSAASTPTF  
PGDAHSQPSFVREQDIAIIGMAFRGPGADDLDAFWNNLVEGVESITFFSEDELLAAGVPREHL  
ASTRYVRAKGELTGMMDFEPEFFGYSAAREAAMDPQFRVFHECSWHALEHGGYDPTRCAASIG  
20 10 VYAGVTNHLPLWLMRTLPHLTEEEQFGALLLTDREFFAPLLSYKVGLRGPALSLQTACSTSLVA  
IGTACREL RAGACQMALAGGVTA SIERCGYFHQEGYILSPDGHTRSFDAAGTVFGDGVGMV  
LLKPLAQALADGDTIHAVIKGIGINNDGARKVGFTAPSRAGQTEAIRAALRDAGVASNRVSYV  
EAHGTATRMGDPIEVEALTQAFRAEADGPLPPGSCLLGSVKS NVGHLNAAAGVAGLVKTVLAL  
25 QHRRLPTSLFYQSPNPHIDFAASEFRVNGQTS DWVAPEGTRLLAGVSSFGIGGTNAHLIVEEA  
15 PKALPTTAAPLSTEPNDLDAGDADGLVLPISARTPTALAH IATNLNHLERHPTIALADVALT  
LQLGRRQWPHRHS LICNRTEAIKLLRAVVHSAEVPPAQAPVSDAPRCVFLFPQGAQYPSMA  
RDLVRNCPDFALHLDPCLDQLAEELLPEDPRC ILFGDGPADRLDQTAYTQPLLESVSYALARWL  
30 GDFGIRPDAMIGHSLGEYVAACLAGLFSLS DALLLVSEGRMLMGSAAARGAMLAVPLPEWELEE  
RLELLADDRISIAAVNTAESCVIAGPSEAIERCAQRWAAOGLTCTPLRTSHAFHSAMMEPIVE  
20 PFGHVLARVTFAPPRARWISNLDGKPIDSAAVMQPDYWVRHLRQPVRFHEGLSHLLAEDTHAW  
35 VEVGPGRTLSSFVRRHPAYRHQPIVNP MRHVESTGDVRRWRQALGELWRAGMPVAWERQRRG  
RHAGRRVPLPGYPFERRPFAARRPVELAQAPKAELVKNPDPARWLYRRVWRPAQAAAGGLAV  
QATVLVFGDGSEL CRAAVAQVQRQGLKCVSITAGRQFARESDMRFTLDPADPRQLDQLFAALD  
GSGSRPRYVLHLLTLNPPPDASAIHAHSYSPMALAHALGAHEIAPV SITVVTAGVVAVADEA  
40 25 IREPLQALIVGPCLVIPQEFPGLSVRLLDVNVDDPAPRLAERLVAELSGTDH MVALRGGERLV  
ADV DQVDGLGVGI AKVPLRREGHYLILGGLDIGYHCARYLAQTYRAKLTLTARSSLPPRASW  
ERMLREGNLD SRQRTRIERVL SLEACGAEVQTA AVDLGDRHRLADV FREARGRFGAIAGVIHS  
45 AGIPGHVHSIDELVRVRDEAQFTAKVRGLHHLAEVVDPLNLD FCLLFSSLSTVLGGLGYGAYA  
AAKAYMDSFARRHDRPDECRWIAVNDAWLFEAKTSSVGAELARLAI VPEDAPALFARVLRL  
30 PQSFIVSTADLRARIDTWIRDKNRVPPAEIRAVQPRPDLSQAYAPP IGPLEIQLCGLVSAYCR



5

10

FDRIGRDDSFFEIGLSSFDLIQLSSRIHRITGKDLNTTQLFSYPTVRALALFLGGEPEGLAAE  
EPAMENLWLQRSDATLDE

**pEPOcos6\_ORF13 sequences:**

5

15

**(1) nucleotide sequence**

Seq ID No 65 (>pEPOcos6\_ORF13.seq)

20

10

ATGAAATACGAAACCACCGATTGGAATTGGCCGTCATCGGTCTCGCTTGCCGCTTTCCAGGC  
TCACCCGATCCCGAACAGTTCTGGTCTGAATCTGCGCGCAGGTCGCTCCGGAATCCGCCATTTC  
AGCGATGCCGAGCTGAGCCACATCCCCGCATCCCTGCGTCACCATCCGCATTACGTCAAGGCC  
AAAGGCGCGCTGGACCACGCCGATTTTGAACCAGCCTTCTTCGGCTACTCGCCCAAAGAGGCC

25

15

GAGGTGATGGACCCTCAATTCCGGCTGCTCCATGAGTGCTGCTGGGAGGCGCTGGAGTCAGGC  
GGCTATGCGCCGAGCCAATTTCGCGGGTCGGATCGGCTTGTTTCGCGGCGGCGGCCTTCAACGAC  
GGATGGATCGCCGGTACCCTCGACCGGTGCGCACCGGCGTGGGTTTGAGCTCCCTGGAAACC  
GCGTTCTTGACCCTGCGCGATTACCTGACCACCCAGATCTCCTATCGGCTCGATCTGCGGGGC  
CCCAGCCTGCTTGTCCTAAACCGCCTGCTCGTCGTCGCTGGTGGCGGTCCAGCTCGCCCAGCAG

30

GCGCTGATCTCCGGCGAATGCGCCCTGGCCTTGGCTGGCGGCGTGTGCGCGACCGATCCGCTG  
CATTCCGGATACCTCTATGAACCCGGCAACATCTACGCGCGGACGGCGTCTGCCGACCGTTC  
GACGAGGCAGGCGCCGGTACGGTCTTCGGCGACGGGTGCGGCATGGTCCTGCTCAAGCGGCTG

35

20

AGCGACGCCCAGCGCGACGGCGATACGATCTGGGCGGTTCATTTCGCGGGGCGGGCGTGAACAAC  
GACGGGCACCACAAGGTTGGCTACACGGCTCCTGGCACGAGGGGCCAGGTGGCTTTGCTTAA  
AGTGTTTATCGCGCGAGCCGGTTCGACCCGGCGACGCTCGGCTACCTGGAGGCCCATGGCACC

40

25

GGCACCGCGCTCGGCGATCCAATCGAGGTGAGGCGCTTACCCAGGCCTTCGCCAGCAAACGT  
CGCGGCACCTGCGGCTTGGGCTCGGTCAAGGGCAACCTGGGTACCTCAACACGGCGGCCGGC  
ATCGCTGGACTGATCAAGGTGGTGCTGGCGCTGAAACATCGCGAAGTGCCACCCACCCTCAAT  
CTCGCCCGTCCCAATCCGAAAATCCGCTTCGACGAGACGCCGTTTTTTCCAGTCGTCGAGTTG  
CAACCCTGGCCAAGCGGGACCGGCCCTTTCGAGCCGGCGTGAGCTCCTTCGGCATCGGCGGT

45

ACGAACGCCCACGTATCCTCGAGGAGGCACCGCCGACGGCCAACCCGGCGCCACACGGCAGA  
TTCCGACTGTTGCCGCTTTCGGCCAAGACACCGGCTGCGCTCGAAGCGAAGCGCCGCGATCTG

30

GCCGGCTTCCTCGAACGCCACCCGGAGACCTCCTTGGCCGACCTCGCCTTTACCCTGCAACGC

50

55



5

10

15

20

25

30

35

40

45

50

55

GGCCGCGAGGTCTTCAGTCACCGCGCCTGCCTCGCCGTGGAGACCTTAACGTCCGCGCGCACG  
CGGCTGAGCGGCGAGTCGTGAGCACTTGCGTGGTGGGCCCCGCGCCAGCGCCATATTTCTG  
TTCCCTGGTCAAGGCAGCCAGCTCGCCGGGATGGGCCGCGGTCTGTATCACCATTTGAGCCG  
TTCCGCACGGCCGTCGATGCCTGTCTGCGCGAGCTGGAGCCAGGACTGCGGCAAGCGCTCAGC  
5 GCCCATTTGATCCGAATCGCGGCGCGGACCCACCCGATTGACGACCTTCGTCCAACCCTTG  
TTGTTCCCTCGTCGAGTACGGGGTGACCGAGTGGCTACGCTGCTTGGGTGTGCGGCCAACAAATG  
15 GTGTTGGGTACAGCTCTGGCGAGTATGCCGCAGCCTGCGTCGCGGGCGTTCTGTGCGCCGTCC  
GCGGCGGTCTCGCTGCTGGCCGAGCGCGAGCGGCTGCTGCGCGACCTGCCAGCCGGCGCCATG  
CTCGGCGTCCCGCTGGCCGCCGAGGCGCTCGAGGCGATGTTGCCCGACGCTCTCGATCTGGCG  
20 GCGATCAACGGCTGTGAGCTTTGCGCCGTGTCCGGGCCGGTCGCGGCGGTCCACGCCTTCAAG  
GCCCAACTGGAAGCCCGCGGACATCACGCCCGCCTGTTGCACACCGATCGCGCCTTCCACTCG  
CGGCTGGTAGCACCGGTGCTTGACCGGTTCCAGGCAGCCGTTCAACACGTGGAGCTGCGGCGG  
25 CCGCAAGTACCTTACCTCTCGACCGTCAGCGGGCGATTGGAGGCGGATGGGCCGGCGAACCCG  
CACTACTGGGTGCGTCACCTGCGCGACACGGTGCGGTTTGGTCCAGCCCTGGAGGCGCTGCCG  
15 CCGGTGGATTCTTCGTGTGCATCGAGGTGGGACCAGGCTCGGCCTTGAGCACCATGGCGCGC  
GAAACGTTGGGTTCAGGCGCGACTGATTTGCTTGCTGCCGCGGCCGCGAACGGGGCAAATC  
30 GAGCCCGGTCCGGTATTGAAACGACTGGCGGCGCTTTGGCGCAGCGGGTTGACATTGGATTGG  
TCTAAATTGACGGGCGGCGAAGAGGGTCATCGAATTCCCTTGCCAGTCTACCCGTTTCAGCGC  
AGCCATCTGTGAGCTCCCTGGCGGCGGGCCACACGCCTTCGTCGCGGCTGCAGTCGAATCA  
20 GGCGCCATCCTTGCCGAGCGATCCGCAGGGGAAAACGCTGAAACCCGGGATTGCCCGCTGCCA  
35 ACCGCCACGCTCGAGCCCAAGGCGGTGCGTCCGGCCCCACTCGAGGCTACCGACGCCGCAGGT  
ACTCGCGAGCGACTGGCCGAACTTTGGCGCGAGTTGCTAGGGTTGACCTCGATTGGGCCCCGAC  
GACCATTTCTTCGACCTGGGCGGCCACTCGCTGACCGCCACGCGGCTGCGCGCCCTGATTAC  
40 CAGCGGTTGATGTGATCTCGGGCTCGACGAAATCTTCGCTCATTCGCGTCTCTCCAGCTG  
25 GCCGCCGTATCGAGGCGGCGGCCAAGAGCCGATTTTCTCCATTCCCAGCGCGCCGGACCAG  
GACGACTATCCCTTGTCATCCGCCAGCAGCGGATTACAGCATCGTCACGAGGGCCGAGGTC  
GGCACTGCTTATAATTTCCGATCGTCCTCGAGCTGCAGGGCGCTCTGGATCGAGTGCGATTG  
45 GAGGCGACGTTGCGGCGATTGTTCCGGCGTCATGAGGGGTTCGCAACCCGCTTTGTGATGCGC  
GATGGCGGGCCGCGCCAGCGCATTGTACCGGACGTGGCGTTTCGCCTGCCGCTACCCAGGTC  
30 GAGCCAGAGCAGGTTCCCGGGCGCATCGAGGCCTTCATCCGTCCCTTCGATTTGGAACGCGCG

5

10

15

20

25

30

35

40

45

50

55

CCGCTGTTCCGCGCGGAGCTGTTGCAGTTGGCCGAGCAGCGCCATCTGCTACTTTTCGACATG  
CACAACCTTAATTGCCGACGGTATCTCGCTCAACCTGTTTCGTCGCCGATTTCGCGGCCCTGTAC  
CATGGTCGTCCGCTGGCGCCGCTGAAACTCCGCTATCGCGACTATGCCGTTTGGCAAGAGGCG  
CGGCTGGCCTCCGATGACCTGCGCAGCCAGCGCAATGGTGGCACCGGCGGCTTTCGCCGCCG  
5 GTGCCACGCTGGCGCTCCCTCCCGATTTCGCGCTCCGGCGGTGCGCCGCTACAAGGGCCGT  
AATGTGGTGTTCACCTGGACCGGGAGATCCGCGACCGCCTGGTGGCCCTGGCTCGAACCCAG  
GGGGTCACCATGAACGTGATGATGCTGGCGCTCTGGGCTGCGCTGCTGCATCGCGAAACCGGC  
CAATCGGAGCTGGTGGTTCGGATCGCTGCTCGGCGGGCGGCCGCACAGCGAGCTGCATCCCGTG  
ATCGGGCTCTTCACCAACTTTTTGCCCCTTGCGGTTGGCGGTTCGAGGGATCGACCCGCTTCGAT  
10 CGCTTCCTTGCCGCTTGCCACCAGGTGTTTCTCGAAGCCTATCAGCGCCAGGACTATCCGTTC  
CACTTGTTAGTCCAGGAACCTCGTGCCGGTCAGGGACCCGTCGCGGTTCGCCGCTGTTCCAGACC  
TCGCTCGTCTACCACAACGAAATTGACGGCAAGACCAAGCTGGAATTGGAAGGGCTGAAAGTC  
GAAGTGGTTCCTTCGAAAAGGGTGTGGCGAGGCTGGATTTGAAGCTGGATGTGACACCTTTT  
25 TCCGACCGACTCGAATGTGTTTTGCAATACGACTTGGATCTGTTCTGCGAGGAGACGATGCGC  
15 GGCCTGATCGCGCGGTTCCAGGCGTTGGTGGCGGGGCTTGTCGCCGATCCGGCGCAATCGCTC  
GCCGCCGCGAGCGTTTCCGGGAAGCGGGCGCTGCGCGCGGGCGTGGCCACGGCAAGCGAATCG  
TCGCCGCGAGTCACTGCCGCCGCAACCATCGACGGCGTACGCCACTCCCTCACCGCAGTCACCG  
TCGCCGGTAGTCCTGACGGGACCCGCCGACCTGCCCGCGATCTTGGCGGCCTACGTGGGGCAG  
AACCCCCATCCGTTTCGCGATCCATCGGGGTCTCATTTTGGAGGCGCCGCTGGGGTTGCGAGCG  
20 CTGCGGTTCGGCGCTGGACGCGAGTGCTCGGAGAACACACCCATTGGCGCAGCGTGCGTGCGGGC  
35 GATCGCGCGCGGCGCGTGGATAAGTTGGAATTGACCAGCCTGGTGGCGGCTCGACGACCTGCGC  
GGGTTGGTCAATCCTCAGGCGAATGCCTTCACCCTGGCTTGGCGCGATCTGGCGATGCCGTTT  
GGGGAGGGGCGTCCCCTGTGGCGACTCCGCCTGGCGTGGTTCGGCTCCATCGCGCTGGTTGCTA  
40 TTGCTGACGGTTCATCCATTGATCGGCGACAACGGCACGGTCGACCTCTTTCTGGCGGCACTC  
25 GCCGATCACCTGCGCCGCGCGTCCGCTTTTCCCGTAGCACCGCTCGATGAGGCCGAGCTGGAG  
GCCGAGCTGAAGTGGGGAGAGGAAGGGGAGGGCCTCGGGCTGACCGCGATCGCGCCGGTCCTG  
GGCCAATTGCGCGAAAGTCGGCTGAGTCCTGTGGCCCAGATGTGGCTGGACGAGGTCTGTGCG  
45 CGCCACGACCTCACCCCGCTAGAGGTCTTGGCGGCCCCGGCTCCTCGATTGGACACGAAGCCAC  
GGTCACGGGTTCGATCGCTTTGTGGACGCCGCTGCCCGAGGACCATCCGCTTCGCGATGAAGGC  
30 CGCTGCCTCCAGGTTTCGCCTGCTGGAGGGGCCCGCGTTCGCGAGGAGCGGGCGATCCAAGC

5

10

TGGCTCGAGCAAATCGCCTTGAGACGGGGTACCCCTGCAACGGAGGTCGTTTGCCCTACTCCG  
ACCCAACGGGCAGCCATCGACCTCGCGCTGGCCTGGCTGCCGCAGCCGCCTCTTCACGGTTTG  
GTCGGAACCGTTCAGCCGTGGCCGGAATCTCCATTGGTCTGTCCGTTTCCCCTCAATCTCGCG  
TTCCGCCCAAGCCATCCAATTGCCTACGCGCTCAAGCACGAGGCCACGCTCGCGGTCACGGCA  
5 CGGGCGCGCGATCTGATGCGTTTCTCGACGGCTTGGGCCCCGAAAGC

15

**(2) peptide sequence**

Seq ID No 66 (&gt;pEPOcos6\_ORF13.pep)

20

10 MKYETTGLELAVIGLACRFPGSPDPEQFWSNLRAGRSGIRHFSDAELSHIPASLRHHPHYVKA  
KGALDHADFEPAFFGYSPKEAEVMDPQFRLHECCWEALESGGYAPSQFAGRIGLFAAAAFND  
GWIAGTLDRLRTGVGLSSLETAFLTLRDYLTQISYRLDLRGPSLLVQTACSSSLVAVQLAQQ  
ALISGECALALAGGVCATDPLHSGYLYEPGNIYARDGVCRPFDEAGAGTVFGDGCGMVLLKRL  
25 SDAQRDGTIWAIVIRGAGVNNDGHHKVGYTAPGTRGQVALLKSVYRASRVDPATLGYLEAHGT  
GTALGDPIEVEALTQAFASKRRGTCGLGSVKGNLGHNLNTAAGIAGLIKVVLALKHREVPPTLN  
15 LRRPNPKIRFDETPFFPVVELQPWPSGTGPLRAGVSSFGIGGTNAHVILEEAPPTANPAPHGR  
FRLLPLSAKTPAALEAKRRDLAGFLERHPETSLADLAFTLQRGREVFSHRACLAVETLTSART  
30 RLSGESSSTCVVGPAPSAIFLFPQGSQLAGMGRGLYHHFEPFRTAVDACLRELEPGLRQALS  
AHFDPNRGADPPDSTTFVQPLLFLVEYGVTWLRCLGVRPTMVLGHSSGEYAAACVAGVLSPS  
AAVSLLAERERLLRDL PAGAMLGVPLAAEAEAMLPDALDLAAINGCQLCAVSGPVAAVHAFK  
20 AQLEAAGHHARLLHTDRAFHSRLVAPVLDRFQAAVQHVELRRPQVPYLSTVSGRLEADGPANP  
35 HYWVRHLRDTVRFGPALEALPPVDSFVCIIEVGPGSALSTMARETLGSQARLISLLPRPRTGQI  
EPGPVFERLAALWRSGLTLDWSKLTGGEEGHRIPLPVYPFQSHLSSSLAAGHTPSSRPAVES  
GAILAERSAGENAETRDCLPTATLEPKAVAPAPLEATDAAGTRERLAEWLWRELLGLTSIGPD  
40 DHFFDLGGHSLTATRLRALIHQRFVDLGLDEIFAHSRLSQLAARIEAAAKSRFSSIPSAPDQ  
25 DDYPLSSAQQRIHSIVTRAIEVGTAYNFPIVLELQALDRVRFEATFAALFRRHEGFRTFRVMR  
DGGPRQRIVPDVAFRLPLTQVEPEQVPGRIEAFIRPFDLERAPLFRAELLQLAEQRHLLLFDM  
HNLIADGISLNLFVADFAALYHGRPLAPLKLRYRDYAVWQEARLASDDLRSQREWWHRLSPP  
45 VATLALPPDFPRPAVRRYKGRNVVFHLDREIRDRLVALARTQGVTMNMMLALWAALLHRETG  
QSELVVGSLGGRPHSELHPVIGLFTNPLRLAVEGSTFRFLAACHQVFLEAYQRQDYPP  
30 HLLVQELVPVRDPSRSPLFQTSLVYHNEIDGKTKLELEGLKVEVVPFEKGVARLDLKLDTVTPF

50

55

5

10

SDRLECVLQYDLDFCEETMRGLIARFQALVAGLVADPAQSLAAASVSGKRALRAGVATASES  
SPQSLPPQPSTAYATPSPOSPSPVVLTPADLPAILAAYVGQNPFPFAIHRGLILEAPLGLRA  
LRSALDAVLGEHETHWRSVRAGDRARRVDKLELTSLVRLDDLRGLVNPQANAFTLAWRDLAMPF  
GEGRPWLRLRLAWSAPSRWLLLLTVHPLIGDNCTVDLFLAALADHLRRASAFPVAPLDEAELE  
5 AELKWGEEGEGGLGTAIAPVLGQLRESRLSPVAQMWLDEVCRRHDLTPLEVLAARLLDWTRSH  
15 GHGSIALWTPLPEDHPLRDEGRCLQVRLLEGPPSQRGAGDPSWLEQIALRRGTPATEVVCPTP  
TQRAAIDLALAWLPQPPLHGLVGTVPWPESPLVCPFPPLNLAFRPSHPPIAYALKHEATLAVTA  
RARDLMRFLDGLGPES

20

10 pEPOcos6\_ORF13.1 sequences:

## (1) nucleotide sequence

25

Seq ID No 67 (&gt;pEPOcos6\_ORF13.1.seq)

ATGACGCAAGCCTCGGCCGCGTACGACGTCCCAGGTCGCGCCGGAGGTCACCCCCGGCCGAAAG  
15 GACGACGATGACGATCAAATCCGAGATGTGCGCCGTTGCTCACTCTGCGGAGACCGGCTTCCG  
CGCTGGGCCACGCGTGGGCGGCGGATGAAGCGGGCCGGACGCCGGAGCAGGCCGGCGTGAA  
30 GCTGCTCCGCGCCCCGGTGAAGCGGAAGTGGCTGCCCCGGCGCCCGTCCTGCGCCTGAGCGA  
GCGGCGTATCCCGGAGGTGTGGGCAGGCTACCGCGCGAGCGCGGGATGACCCGAGCCCCGCCC  
GCCGGCGCGACCATGACGCCGCCCCACGGGGCGAGTCGTCCGGCGCGCCGGCGCGCGTCCGGG  
20 CTTCCGCGCCCGGGCGGGCAGGTGCAGGATGGTCGGGCATGG

35

## (2) peptide sequence

Seq ID No 68 (&gt;pEPOcos6\_ORF13.1.pep)

40

MTQASAASTSQVAPEVTPGRKDDDDQIRDVGRCSLCGERLPRWATRGRRDEAGPDAGAGRRE  
25 AAPRPGEAEVAAPGARPAPEAAYPGGVGRLPRERGMTRAPPAGATMTPPHGASRPARRRASG  
LPPPGGQVQDGRAW

45

50

55

5

## pEPOcos6\_ORF14 sequences:

10

## (1) nucleotide sequence

Seq ID No 69 (&gt;pEPOcos6\_ORF14.seq)

5 ATGGTGACGCGTCCGACGTCCGACGGCATCGAGGACGAGCTCGCGCCGTCCCCCGGTCTCTG  
CGCGGCTGGCTCATCGAGGGCGAGCTCGGCCGCGGCGGGATGGGGCGGGTGTTCGGGCGCGG  
15 CACCCGAAGACGCGGGCGCGGGCGGCGATCAAGGTGCTGCTCGGCGACTACGCCCCGCGGCCG  
GACGTGGTGGCCCCGCTTCGGGCAGGAGGCGATCGCCGTCAACATCATCAACCACCGGGAATC  
GTCCGCGTCTTCGACTCCGGCGAGCTCGAGGACGGCTCGCCCTACATCGTGATGGAGTACCTG  
20 GACGGCCGGGGGCTGCGCGACTGGGTGCAGGCCGTGCCGCCGCGGAGCGGCCGCGGCAGGTC  
GTCCGGCTCGGCTACCAGATCGCCTCGGCCATGGCCGCGGCGCACGCGTCCAAGGTCTCCAC  
CGCGATCTGAAGCCGGAGAACATCATGGTGGTTCGAGGACGAGCTCGCGCCCCGGGGGAGCCGC  
GTCAAGATCCTCGATTTTCGGCATCGCGAAGGTCCTCTGGGGAGGTCTGCCCGAGGTGCTGGAG  
25 CTCGAGGGGAGAGGCTCCCTCGCGCCCCGCGTCCGCGTCCACGATCCGCACCGAGCTCTCGACG  
CGGCCGGCGCCGACGGTGGGCGCCACGACCGGCCAGAGAGCCCGCTGGGCGCGAGCGCCACG  
15 CCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCA  
GAGAGCGAGGCCACGAGGAAGACGCGCTCCGGAGCCTCCCCGTCTGACCGAGCGGCAGGCCC  
30 GCGATCCACCCCGCGCCGGTTCGAGATCCCGCCCCGAGGCGGTCTCCTCCGCGGCGTCCGCGGGG  
TCGCGCGCGTTCGATCGAGCCAGGCGCGCCCCGCGCCGAGAGCGAGGGCGCGGGACAGCCACG  
20 ATGCCGTTACGCAAGAGGGCGTGTGGGGCCTCGGGACGAGGAGCTACATGGCGCCGGAGCAG  
35 GAGCGCCACTCCGGGAGCGTGGACGTGAAGGCGGATGTCTACTCGCTCGGCGTCATCCTCTAT  
GAGCTGCTCGAGGGGCGGACGCCCCGACGCGCCGAGCGCCGCGTGGCCGCCCCCGATGAGCGCC  
GCCACGCGCCCGATCTCGTCCGCCCTCGTCCACCGGGTTCTGGCGTTCGATCCCGATGCGCGG  
40 CCGCGCATGGCGGAGGTGGCGAGCGCGCTTACCGGCTCGGCCGGGCGAAGAAGGAGCTCGAC  
25 GAGGCGCTCTCGAGGTGGGTCTGTCGGCGAGGGGCGCCGGGGCTCTTGCCGTGCGGCTATGCT  
CTTCTCGAACTGGTCCTCCTGGGCCCTGGGAACCTTATACGATTCTTTCCAGCCTGTAAGTGCA  
TTTTTCTTTCAATATCGTCCTCTCTTCATATACGAGGTGAGTTCTCTGAGGTCTCCTATAAG  
45 TCTGGGGTGTCTATTTCGGCCTCTTACTTGTTACTTCGCCTTCTTAGGAGTTTTTCCTTAATT  
TTGCCCTCTTACATTCCCGTATTCACTTAACTGGGCCCTATCTCATTCGC

30

50

55

5

## (2) peptide sequence

10

Seq ID No 70 (&gt;pEPOcos6\_ORF14.pep)

MVTRPTSDGIEDELAPFPPVLRGWLIEGELGRGGMGRVFRARHPKTRARAAIKVLLGDYARRP  
DVVARFRQEAI AVNIINH PGIVRVFDSGELEDGSPYIVMEYLDGRGLRDWVQAVPPAERPRQV

5 VRLGYQIASAMAAAHASKVVHRDLKPENIMVVEDELAPGGSRVKILDFGI AKVLWGGLPEVLE

15

LEGRGSLAPASASTIRTELSTRPAPT VGATTGPESPLGASATPESALGASATPESALGASATP  
ESEAHEEDALRSLPVVTSGRPAIHPAPVEIPPEAVSSAASRGSRASIEPGAPAPQSEGAGQPT  
MPFTQEGVWGLGTRSYMAPEQERHSGSVDVKADVSLGVILYELLEGRTPDAPSAAWPPPMSEA  
ATPPDLVALVHRVLA FPDARPRMAEVASALHRLGRAKKELDEALSRWVVG GAGPGLLPCGYA

20

10 LLELVLLGPGNLYDSFQPVSAFF FQYRPLFIYEVS SLRSSYKSGVSYAS YLLLRLLSFSLI  
LPSYIPVFILTGPYLIR,

25

or DNA sequences complementary to said open reading frames,

- 15 (b) DNA-sequences which hybridise under stringent conditions  
to regions of DNA sequences according to (a) encoding proteins  
or to fragments of said DNA sequences,

30

- (c) DNA-sequences which hybridise to the DNA-sequences accord-  
20 ing to (a) and (b) because of a degeneration of the genetic  
code,

35

- (d) allele variations and mutants resulting by substitution,  
insertion or deletion of nucleotides or inversion of nucleotide  
40 segments of DNA-sequences according to (a) to (c), wherein the  
25 variations and mutants offer isofunctional expression products.

45

10. DNA sequence according to any of claims 1 to 5, wherein  
the DNA is selected from the group consisting of

30

50

55

(a) the following DNA Sequence:

Seq ID No 71 (>Contig43)

CGGGTATTTGTGATATGTGGGCNGTAGTCGTATGCTTCATTAAGTACATC  
CGTCCGTNGTAGAGAGTGACTCTGTGCGCAGCGATAATAGACACGCTTGTG  
ATGCTATAGGGAACATAGAGTCNTAGTAGATGATACGACGAGATATTNGT  
ATAGAGCGTATAGACCGACGTGTGAGCGTCATAAGTGTTGTGTGTCATGA  
GTGTGCTCAGAGGACGTGCAGACATTATATGAGCAGATGATGAGAGAGAA  
TCAATGCTGCAAGNTATTCGTGCAATCTACATTATATCGAATCGTGTATG  
TGCSTTTGTGCGCAGCGCGATNCGATGAGATACCGAAAGGGTATGTATCTA  
TNTTCGTGACGCTCGATNAGAGCAAATCCGCTACCGTGGAGATATCGTGT  
ATCGACTCCATCACGATCAGTATCATGATACGTCAAACGAGTACACTCAT  
TATTGATAACACACGTANGTGTGCATGCACAGTTATCGAGTGTATTGTGT  
GCATGAGAGGTATAGGATNTATAGGCGAGCATATATATCTATATATATAG  
GTTAAGAGTAGAANACTATGAAGATGCAGGAAGTAGTATCTCGCGGACAA  
ACGGNGTACCTAGCGGGGTTGAAGTATTATCGACAGTGTATAACGACTCA  
ACAGGGNTACGAGGTACATTGTATTTACAGTGGTTGGAAGGATTGCGCGA  
GGAAAGGTAGTGGTACCGTGTGAGCTACGATGCTCGGGATAATGGTGATT  
AGATAGAACCTTAGCGTTGCTAGATGAGTGAGTGGTGGTATGAGTAGAGT  
TTTTGTTCTAGCTTTGTGTCCAGCGAGGATTCGTTTCAGTCTGAAGGGTAA  
GAGTACGTCCATCGCACACCCGACCGTTTTGAGGAGTTCTCGGTGCGTGG  
TCASTGGGGTTTGGAGAAGACAGAGTTGATTCATAGGGTTATCAAACGAG  
TTATGTGGATAGATGGTAGTGACCCCATTTGAGTGAGAGTGTGCGGTTA  
ACANCAGCAGGATNTAT

SEQ ID No 72 (>Contig44)

TAGGTCTTTGACACCATGGGAGCTGCTACCGATGTTGCCGAGCACGATCG  
CGCCGGCGCCGACGAGCGACTGCAAGCCGGCCGCGGCCATTTACGCCTGA  
CGAGCGAGGTGGGCGAAGTGCTGGTGCGCGCCGTGCGTGTGAGCGCGCC  
CAGGTCCGCCGTTGCGCCGTGCGCGAGCAGTAGCGCGCCGTGCAAGACGA

5

10

15

20

25

30

35

TCACCGCGATCGAGGTCAGCGTCGTGGGGGCGAGGCCGAGGAGCGCGAGG  
ATGCCGAGCACGACGCCGGCCGCGCCGTAGACGAGCTTGGCGCCCATGCC  
GCCGCCGAGCTCGGTGCGCGTGTCCCACTCGACGGGCGGCGCGGTGCGGC  
TCAACGCGCCGAAGCGCGAGGCGATCGCGCCGCCCTGCGCGATCAGCGCG  
5 GCGCCGAATACGATCGTGGCGATCTGAGGTGAGCTCTACTGGCATGATCC  
CCGTCAGCCCCGAGGATGGTGAGGACAATCGTCGCGGCGCCGCACAGTACC  
TCGCACGAGCGAGCCTCCGAGCACGACCTTCGGCGTCGTCTCGTCCTTTG  
GTCTGCGTCGCGCGCCCGAGTGCGGCGTTATGTGGCTCTCCGGCTGTGCA  
AACCGTTCACGTTCTTCCGGTCTTGAGTCAGCATCGGCATGATTCCCCC  
10 GTCTGCGGTGAGGCCTTGTGCGCTCACGCGCGCTCCGACTTGACGCTG  
CTGTGCCGGGTTCTCTCGCTCAGGAGGCGCCTCTCTTGGTGGTGCTTGCG  
TCCTGGTCCGTTTGGCCGCCTGTGCGGTAGGTTTCTTGAACCAGGTGACC  
TTCAGGGACCCCTTGATGCGCTCCATCGTGTCTATGTCGATCCTTCTCT  
GACTTGTATGGGTCTCGAACC AACTACGCTTGATCAGGCCTTCGAAGGGT  
15 CCTTTGGGAGATCGACTCTGGATCCATACCGGGAGCCCCTGTTCTGCCGC  
TCTCTTAAGTTTCCCCTTCTGTATCCGTGTCGACCGGAAACGCTTTATCT  
CTAATGCGCTCTAATTGCGTCTCTGCCACACGTGCGCTTCACTCTGGATC  
TACTTCTTCTCCCTAGTCTTCTACCTCCGTACCCTTATTTGTTGGTTCTA  
TTTATTTCTTTTCGCTTCACCTCGCGTCATTGTGCGCTAGTGTTCTCTCC  
20 TCATATCGCCTTTGGTCTCCCTCGAGCGTACAGTCCTCTCTCTTCAGATG  
CTTTCCGGCTCCTCTTCTGCTGGCCCCCTTATCCTTTCTAATACTTC

SEQ ID No 73 (&gt;Contig48)

40

25

45

30

50

55

ATGCGCCAGGAACACCCCGGTGCGGCTGCCGTCGAGGGACTGGGGTGCG  
ATGCCGGCGTCCTCGAGCCCTTCCCAGGTGACCTCCAGCAGCAGGCGTTG  
CTGAGGATCGAGCGACCGCGCCTCCCGAGGCGAGGTGCCAAAGAACGCGG  
CGTCGAAGCCGTCCACCGCCTCGGTGAGCAGTCCGGCCCAGCGCGGCACC  
TCCTCGCTGGGATGGACGCCGACCAGCGCCCAGCGCCGGTCGAGCGGCTG  
GACCGCGTCTCGGCCTGAGTCGAGCAGCTCCAGAATGCCTCCGGAGTGT  
30 CCGCTCCGCCGGGGAAGCGGCAGCCAATGCCTACGATGGCGATCGGCTCG



5

10

5

15

20

10

25

15

30

20

35

40

25

45

30

50

55

GTCCGCTCTTGCTCCAAAGACGCGTTCTTTTTTCGCAAGCTTGTCCATGAG  
CAGAAGGGCATGCTCAAGCTTCCCGGCATTTCGTGGTCGCCATACTCCCTC  
GGTCCCTTACTCACCAACGATCTGCGCGAGCTGCGCCAGCTTTTCGGCGA  
GCAACGCGTCCTTCTGCTCGTCCGTCATGCCCCGCAGAGCCTCGAGATCT  
GCGGCATCGTTCTCGAAGCTCTTCTCCCGCTCGGTGGCCGGAGCGTGGGT  
CGCGCCGGCATTTCGAAACAGAATGTCTAGCAAGCTCCCGCTCAGAGCTG  
CTACGTTAGGGTAGGTCCATAGCAGGGTCGCCGGCACGGTGATGCCGAGC  
GCGGCCTCGATGCGGTTGCGGAGCTCCAGGCCTATCAGCGAGTCCATGCC  
GAGATTGCTGAACGGCACGTGCCGCTCGATCCTCTCCGGCGGAAGGCGCA  
GCCCCCGCCCCAACAGCTCGCTCAAGTGCTTCTCCAGAATCAACTGACGA  
TCTTCGGGCCTGGCGCTCTGCAGCGCCTCGCGCAGGTTTCGACGCGTTCGA  
CGCGCCTCGGTTCGGCGCGGTACGCTCCTTCAGCAGCTCCGCCCACAGCG  
CCAATCGGGCCGCGTTGGGATAGAACTC

15 SEQ ID No 74 (>Contig49)

ACCACCGCTTCACTCAGTATGTACTTTGTTATACTCGTCTTAGTACAATG  
ATATAATACTCATGTGTATTCTTAATCTCGGGGAGANAAAATTGGAATAC  
TGGACACCGTTGCCGCATGCNGACTCTAGAGATCCCCCTGCGACGGTATC  
CCACGGCACCGGTATGGCCGGCGCGCTCCGGGGGTCAACGCCCCGTGG  
TTGCCTTCACGACAACGCCGGTCGGGCGGGGCGCCGTTTCGATGCCGCGGG  
CCCGCGCGCGGCGGCGGCTTATCCTGTGGAGCATCTGGAGGGCGCTCACG  
CACCTGTCAGTCTAGTTCTGGCCCGCCCGGAAGGAGTCCGGGAGGCCGAA  
GTTGAACCCGATGTAGAGCGCGATGAACGACGGGAGCACGCGCGCGGGGA  
TGTGCAGCGCGGCGCCGATCGGCGTCGCGAACAGGACGAGCTCGCCCGGC  
ATGCCGGGCACGACATACCCGAGCAGAAACACGATCGGCACCACGAGCGT  
GAGCTCGAGCAGCGATATTTTCATGACCGACCGCGGGCGGCGGCGCCCG  
CCATGACGAACACGCAGATCAACGTGCCGTTGACGTTGAGCCAACCGCCG  
AGTCCCACCACGAAGAGCCTGAGCTCCTGCGGCACCGCCGGATAACATTT  
GCGGACGAGGTGCAGGTTGAGCGGCGTCGCCAGCGCCTCGCTGCACGAGG  
CCCACAGCAGCGGATAGACCTTGAGCCAGTAGTTGACGAAATAGTCGCC

5

10

15

20

25

30

35

40

45

50

55

AGCGAGAACTCCGGGGCGGCGGCCCTTCATCCGCAGCAGGCTCGCCGCATG  
GAAGACGAGGCAGGCAGCGCCGACGACCCCGGACACGAGCAGATAGGAGA  
GCATGAGGTCTCCGGCGCCGAGCGAGGGCGCCCCCGCCGAGGCGCGCGCG  
AGGTCCCCCCCCGTGCACCTGCGCGGCGAGCTGCGCGGGCAGCCCGCGGAG  
5 ATAGGCGCCGAGCCCGAACATGAAGAGCGGGACCAGGCACTGCACGGCGC  
CTCCCGCGCGCTCCAGCGCGTCCGCGCGCGCTCCAGCGCGCGGGCGACC  
CGCGGCGCGCGTACGGCCGCGAACGACGTACGATGCCGCGTAAAGGGC  
GAGGAAGCACGGGCTCGAGATGACCAGGCCCGAGGCGCTATAGAGGGTGC  
GCGCGGCCCTCGAACGGCGCGCCCGTCTGTGGCTGGGGAGCAGCGGCAGC  
10 CCGAACACGAGCCATGTGACGACGACCCCGAACAGGCACGCTGCCAGACG  
CTTGAGGGCGAGCCAGCCCATGATGTACGCGAGCAGCCGCCCCGGGCGCC  
CTTGCCGGTGCAGGCTCACGAAGGTGGGCACGAGGACGACGAAGATGACG  
ACCGGCGCCAGCGTGGTGTACCAATGCAGGAGACCGTCCATGGCGCGGGT  
CGACCACCGCGTGACGCTGGTCTCTCTGTCTGACTCGATCATGGCCCAT  
15 CGCCTAAACTAATGATCCGTTCTCAAATTGGTCAAAAAAAGTTCCTT  
AAGACTGTTTTACTCCGGAATATTAATATATTTCTGAGTGTGAGGTGATG  
TTAATCACACATTCTGATATTCTCAAGGGGAATCCGTGTCATTGTGAATA  
CTTCTCTCTCTACAAGAGAGGTTATATATGGTCTCGAATATCTCGTCCGC  
TCTTATATATATTCTCTTGTGATAATATATATCGAGTGTGGGTACTCAGC  
20 TCTCTTGGTGTAACTCTATAACTCGGCATCTCTCATAATACCTTATATATA  
CACACTCTCTCGGTCATATCTCGCAATAAGATATATTTTATATGTTCCG  
CGTTTTATCCGAGTGGGATACACTTTTTCTATATTTTCTTTGGTGTGACG  
CGTGGCGTCCGAGCCTTATTATTGATTGGTAGTCACGATATTCTCTAGAT  
GACATCATACAGATGCTCATAACTCGATAAACACAGGTCGTACACGACGA  
40 GACTCTCACTCTCACTCTT  
25

SEQ ID No 75 (&gt;Contig50

TCCCCAGTTTCTCCTCTCTACGCNCACATCTCAGCAGGAAAAANATAAT  
GGAGAATCGTTGCGCTCTAGCAGCATCTATAGGATCCCCGCTGCTCTTCT  
30 TCATGCACCTCGTGGAGCAGAAGTTCATCAACGCCTTCGCGATCATCGTG

5

10

15

20

25

30

35

40

45

50

55

GCGGTGAGCTTCCTGGCGTTGCTCCTGTCGCTCGTCGTCGCCGACGTCCG  
GACGCGGAACACGTTCCCGCCCCGCGCCTTTGCCGGCGCTGAGCCCCCGG  
CGCCGGCGCTGATCCCGCCGGTGCCGGGCGGATCCGTTGGGCGCTCGCCG  
GAGCCGCTGTCCGTCCGCCGGTGATCCGTTGTGCGGGCGCCGTGCCTCGG  
5 GCTTACTACCCCTCTCGCGGGTGGGGATATGGCCGTGGATGAGGGAGGC  
GATGAAAATCGTGATCGCCACGTGCGCGTTGTTCTAGATCGTCCCAGGCT  
15 GACCGTCGGGAGCGCCCAGCACGAGATGAAGAGCCACACCGCGAGGACCG  
TGTTGAGGTACCGCACCGCAGGGGCGAGCATGGCGGTGATCGCGAAGATC  
ATGCAGAGCAGCCCCGAGCACCCATGTGTTCGTCCGCTGCGCGTGGCTGTG  
20 CGGCCAGATGACGGCCGAGATGAGGAGCCAGAACCCGAGGACGACGTTCA  
CGATGCGCGCCATGAGATTGCCAGCTCGAACCATGCTCCCTCCCACCTCC  
GATCATGGGACCGATCGGGTCGCCACGGATCGATAACGGGCGTCAGGAGA  
CCGTCAATCGGCGAGCTCGTGAGCCATGGCGACAGCCCCCGGACCGCGCC  
25 GGCGGGTCTCTGGCCTGCTGGTCGCCGTGCCGGCGGCGGCGATGGGCCTG  
CCTCCGGTCGGGCCGCGCGGGCGCGCGGCGCTCCGCGCGAAGCTGGAGAC  
15 GCGGGACATCGCTCCCTCGCCCCGCGCTCGGACGAGCGGCGCGAGCCACT  
TCTCGACGGCCGAGCGGGCACTAAGCTTCCGTCATGAGGCTCGGCGCACG  
GCTCACCACGCACACGTTCTCGGCCGGCGCCCGGCATCAGCTTCGTCTG  
TCCAGCCGATCCCGGGCTCGGACCAGCTGTTCTGTCATTCCGATCCAGTAC  
20 CTGCTCGCGGCGTCGCTCGCGAAGGAGCGAGGCGCGCCGCTCTCGAAGGC  
35 GGCCTGGTCCCAGGTCCACCAGCTCATCTGGGGCGGCGGCGCGCTTCGCC  
TCATGCTCGGCTTGACCCTAGGGCTGATCCCGCTGGCCGGCGCGTTACG  
AACGCGATGACGGCGTTCCTCACGACCGAATATCTCGGGTACTACGTGGA  
TAGAGCCCTCGACAACCCGGACAATCCGCCTCCGGCCCTGTGATCCAGG  
40 ATGCTTTGGACGCCATCACCTCCTTCTTCACCGGGCGAGCGCGGTAGGCG  
25 AGC3GTCCCTGGGTGAGCCCCACCCTGCGGCTCTAGGAGCCGAAGGGCGA  
GCTCCTCGGGAGCGGCGGGCGTCAACCACAGATTGCGCCGGCGCTTGCGG  
CC3AGCGTATCGCGACCGCCGCCACCGCCGCCACGGCGAGCACGGTGAC  
45 CGCGCGGCGCCGCGATGGCGACGCTCCGGGCGGTGTGCTCGGCCTCGC  
30 CCTTGATGCCGCCCCACCTCGCCTTGACCGTGACCAGGTGCGCGCGGAGC

5

10

15

20

25

30

35

40

45

50

55

CGCTCCTTGCTGTGCTCGATCTGCTCGGCAGGCCCCGCGGGCGCCACGAA  
ACCTGCGCCGGCGCTGACGTGACCTCGCTCGCTGAGGGGACTGGCGACCC  
TCTCCGTGTGCAACCGGATTGCGCTGGATGGATCCATATGTCCCGCGCTG  
CAATCGTTGCTCCCCGCCGGCACATCGGAGTGCTCGCCGGATCGCGCGGC  
5 AGCGCCGACGCCGTACTTCCATAGGATAGCCCACCCCATCGGACAAGCCG  
GCTCCTGACGGCGGGCACCGAATGTTCCGCCAGACGGGCACAAGGCGCACG  
CCGCGGACGGATCGGCCGCACTGGCACTCCAGAGCGCATCGACGGATGGC  
CGACGGATGTGCAATGAGGCGCCCGCACGAAGCGAATTGTCCCGAATACA  
GCGAAGAAATCTATAGCGATGCGAGCAGAAGGATATGTCTATGGGGGGCA  
10 GTCAGAAACTGGGGACAGTCAACGCACATATTCTCTCCAANTGCTAACGA  
CAGCGTGCGCAGAGGAAGTATCCTACTAGTGTAAGAGGGACATTGATGC  
GACCGCATAAACATTGAGTCTACAACGCGTGAGAGGATGGAACACCCCGC  
CCCTCTGAAGGCTAGACAACCATGAATATGTGCAGAGGAAACACAGAATT  
15 CCAAAGGTGAGAACATATGTAGGATCGCGCCACCCGAGATTGAGTGAAGA  
TATACATATATACTTATATGGATCTACAACATGGCGAACCGAACGTAGCA  
NAATAGTAGATATAATTGTAATACTGAGCTACCGACAGAAAGATACACAC  
GAGTGTACACACATCACACGCAGAGTGGTACCAAATTCACACCATGCGAG  
20 CCACAATGTGACACGGAGGAGCACAGCATGGGCGCCACTATGGAGGAGAA  
ACTACTGCAACCCACATCTGATGGACTGACCGCACGGACGGGACGTGTCT  
ATACATACAGATACATCNGATGGAGGAAGATGCATGTGCGATGATATCAT  
25 CGTCGCAAACTCATATGTGCAAGAAGATATGNGTCAACTCAGCACTACTC  
ACACGATACGTGAACAGGAGTGAAGTACTAGGACATCNCATGGTGTGTCGGCGC  
GTGCACGTGATATCAAACCTCTCTGATCAACCACACACTATATAAGGAGTA  
TCGAGCGGCGATGGAACACCCCCTCACAGCATACGTATATGCACAACGTC  
30 TGAACACTCTNGAGACACAGTGAAGG

SEQ ID No 76 (&gt;Contig51)

GATCCAGTTACGCCCCGCCGCTCGGTACGCGGGGTTTTCGGCGTGA  
CCGCGGACGGTCCGGGCGACAACCGGGGGTTGTCGTCAGCGGTTGCGGTG  
30 GATGTGCGCGACGAAGTTGTCCGCCCGGAGTCGTCCTGCGTGCCGCCGC

5

10

15

20

25

30

35

40

45

50

55

GCTGCCGTACGAGCAGCCCGAGGCACCGCGCGAGCTCCTGCCGCTTCCGA  
TGC GCGGCGGCCTGCTCGGCCTCGTGCTCCGCGATTGCGCGCTGGCGCCT  
GCTCTCCTCGGCGCGTTCCCTGCGTCCGCGCGAACAGCACCTCCTGCACCT  
CCGCGAGTTCCTCGGCGGTAGCGACCCGCACGTGATCCGCACGTCTCCG  
5 AGACCGGTGACGACGTCAACCGCGGAGTAGTACGCCTCCGCGCGGACGAC  
GTCGTCGAGGTGAGTTTACGCGTCGGCTCCGACGACCAGGCGCATCGCGT  
15 CCGAGGTGAGGACGACGCCACCGTCGAGCGGCGCGGCCTCCGGCTCGTAG  
ACGTACTGGAGGTGCGTCTCCTCGTCGTCGTCACCGTCGAGGTGACGTG  
CCGCCGAGGCCTCGGGTCCACTCGATCGCGCGTCGTCCGGCGAGGGCTT  
20 CCTCGTACTGGGCCCCACCACGCGCGCAGCTGCTTCGGCGTGCCGTAGCCC  
TCGGCCATGTGCGGGTTCGAGCCCCGGCGACCTCGATGTCCCACAGTCGGTA  
GAGGATCTGGAACGGCGTCATGGACTTCCGGCCCCGGCCGTCTTGAGT  
CCAGACGGGCGGTCTCCATCGCAGCTGCGCCGGCGGCTTCGAGGTCTTGG  
25 TCGACGGAGTCGGGCCGCTCTCGCTTCCCGTCTTGGTTCTTGGTGAGGTA  
CTCGATCAGCGCGACGTGTCGTCAGCTGACCGGACGATCGAGACCATCACGC  
15 CGTGGCCCTTGCCCTTGCACTTGACGCCGGGGGTGTCGCAGTCGGTCGAG  
GGCTCGAACTTGGGGTCAGCCCGCTTGAGGGCGCCCGCCACATCTCCCG  
GAGCCAGTCTCCAGTCCCCCAGGTCCGTCTCGGAGGGCTCAAAGTGTC  
CGACGACGTACCCCTTGCGCGGGGTGCCAGAGAGCTCGCCGCCGAGGAAG  
20 ACCAGCAGGTTGAGGTGGGGGTGGTAACCGTTCTTCTTGGACCGGGTGAC  
35 CTCAGCCGCGCGGACCATGCCGATGTAGCCGATCCGGTGGCGGATGCCGT  
CCTCAGCGGGACGGACGTACTGCGTTCCGTCTTCCGGGTGCGGCGGGCC  
TCAGGGCGGCGGTAGAAGGCCGGGGCCGTGAGCATCCGCTGGTAGGCACC  
GGGCGCGCGGCGGGGCTTGCCCCGACCGGTGAGGACCGGGGCGCCCTTGT  
40 CGTCCAGGAGAGGCCCCGCCCCAGAGCGCGGCGACCAAGCTGTCGAGGTG  
25 GTGGTCTGGTTATGCCGGGCGGTGAGGACGACAACGGCGAGCGTGCCGCC  
GGCGGCGAGGTGCCGAGAGCACCGGTCTTGATCTCCTCGGTCCGGCCAC  
GGCGGATCGCGGAGGAGCACTCCGGGCATAACCAGATCCGCCCCGAGCGG  
45 ACCAGGCCGATCGTGACGACGTACCCACGGCTCGACTTCGCGTAGATCAC  
30 GCCGGTGTCCGGGTGAGGACCCGCGCCCGCACCCCGCGCAGGCGTCGA

5

10

5

15

20

10

25

15

TCCCGGAGACCCGGTTGAGCACCTTGCGGCCCTGGTAGCGGCGTACGGCA  
GCGGTCGCCGCGCGCCTTGTGGTCTGTTCCGAAAGGGCTGCCGCCCTCTC  
GGACTCTCCCGTTCCTCCACGACTGCCACTTCCGCAAAGTCGCTGGTCA  
GTGGGGGGTGGGAAACTCTGTCAACCCCTTTACCTAGGCGTCCCTTTTTG  
CCAGGGGCGGTCTCACGGGCGGCCTCGGCGGCTCGGTCCGCGGCCTTCCG  
GGCCCCGCGCGGCCCGCTTCTTGACGCCTCGGAGCAGAACTCCTTGGCTC  
GCTTGCCGGGGGTGATCGTGAGGGCGGCGCCGAGCGGCAACGGGGGCCG  
GCGGGGACGCGGGCGGGCGACTGAGTCGGCGCGCCGATCAAAGAGGGGGT  
TGCGGACGCCAAAGCGTCCCTTACGCTGGACACAGACGAGTACCTTGGTT  
GGTAGCCGGGTGGACGTCAGAAGCGGTCAGGGATTAGGACCCCTGGCCGT  
TTCGCTTTTTCTGGAGTTGTTCCGGGTAGATCCTGCCGCATCGCCCGCCTC  
ACGCGTGGCTCGCCGCGCGGATGCCTCAGAGGCCCCACCGGTCTGTCAGGA  
CGCAGACGTCGGCGTGCTCCTGGTGGTGAGTCACCAGCTCGACCACACGG  
GCGCGGCCGCGACGTCTC

30

20

35

40

25

45

30

SEQ ID No 77 (>Contig52)

CGGGATCTGGCCTTCATTAACCAACGACGGGGCAAACATAATAGGCTGGG  
CATTGCGCTTCAGCTCACCACAGCCCGTTTTCTGGGAACATTTCTGACGG  
ATTTAACTCAGGTTCTGCCTGGTGTTCAACATTTTGTGCGGGTACAGCTT  
AATATCCACCGTCCAGAAGTTCTCTCCCGCTATGCTGAACGGGACACTAC  
CCTTAGAGAACATACTGCATTAATTAAGGAATATTACGGCTATCATGAAT  
TTGGTGATTTTCCATGGTCTTTCCGCCTGAAGCGTCTGCTATATACCCGG  
GCGTGGCTCAGTAATGAGCGACCGGGTCTGATGTTTGATTTTGCCACTGC  
ATGGTTGCTTCAAATAAGGTATTACTGCCCGGAGCAACCACACTAGTAC  
GTCTCATCAGTGAAATTCGTGAAAGGGCAAATCAGCGGCTGTGGAAAAAG  
CTGGCCGCACTGCCGAACAAATGGCAGGCAGCTCAAGTGATGGAGCTTCT  
GGTCATTCGGGAAGGTCAGCGTGATCAGCACTGGAACAGTXXXXXXXXXX  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGCTGGAACGATAT  
ATCCGATTACGAAGTCTTGAGTTTCCCGACTGAACTTTTCCGGTCTGCC  
TGCCATTCAACTGCGTAATCTGGCTCGTTATGCTGGCATGGCGTCGGTAA

50

55

150

5

10

5

15

20

10

25

15

30

20

35

40

25

AATATATCGCTCGAATGCCACAGCAGAGAAAGCTTGCTGTACTTACTGCA  
TTCGTTAAAGCACAGGAAATAACGGCATTAGACGATGCCGTTGATGTGCT  
TGATATGCTAATTCTGGACATTATCCGCGAAGCAAAGAAAACCGGGCAAA  
AAAAAAGACTCAGGACACTGAAAGATCTTGATCAGGCCGCATTGTTACTG  
GCGCGGGCATGTGCATTGTTGCTGGATGATAATACAGATGTCCCAGATCT  
CAGGCAGGTTATCTTCAAGTTCGTACCCAAAAACAGACTGGCAGAATCTG  
TAAGCAAGGTTAATGAACTTGCTCGTCCACAGAACAXXXXXXXXXXXXXX  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXAAACGTTTTCTTCCGGCGGTGTT  
GCGGGACCTGCATTTCCGTGCGGCACCGGCAGGTGAACATGTACTGGCTG  
CGATTCATTATCTGGCAGAACTGAATGGTTCGAAAAAGCGCATCCTTGAT  
GATGCGCCTGAACATATTATCACCGGTCCCTGGAAACGCCTCGTATACGA  
TGCGGAGGGACGGATACAGCGTGCAGGTTATTTACTATGTTTGCTGGAAC  
GCCTTCAGGATGCACTGCGCCGCCGGGACATCTGGCTTGAAAACAGTGAT  
CGCTGGGGAGATCCTCGCGAGAAGTTGTTGCAAGGTGAAGAGTGGCAGAC  
TCAGCGTATTCCTGTCTGTGCGGGCACTGGGACATCCTGTGATGGACGTA  
AAGGTGTGCAACAACTGGCTATTTCAGCTGGATGAGACCTGGAAAGCCGTG  
GCATCACGATTTGAAAAGAATGCGGAAGTTCATATCTGTAATGAAGGTAA  
ATATCCATCCCTGACTATCAGTTGTCTGGAGAAACAGGAAGAGCCACCAT  
CATTGCTTCGTCTAAATAATCGGATCAAACAGCTACTCCCACCGGTAGAT  
TTAACGGAAGTGTACTTGAGATAGATGCCCAGACAGGATTTACACATGA  
GTTTGCGCATGTGAGAGAATCTGGTGCTCGAGCGCAAGATTTGCACATCA  
GTTTATGTGCGGTATGAATGGCTAAGCCCTGTAATATGGGCCTGAACCCG  
TTGATAAAGCACAATATACCAGCATTGACCCGCCATCGGCTCAGTTGGGT  
GAAACAGAATTACCTTCGTGCAGAAACGCTGGT

SEQ ID No 78 (&gt;Contig53)

45

30

ATTCCACGCGCTCACGGTCAGCTTCGACCCGCGCGAGCGCCCGGCGGCCG  
CCTCGCAGAAGCGCGCGGTACGCTGTCCGAGCTCGGCGCGGACGCGCAG  
GCGCCGGAGTGGCCGTTCCCTCGTCGGCGACGAGGCGGCGACCCGCGCGCT  
CGCCGAGGACCTCGGGTTCGCTACGCCTACGATCCGACCACCGATCAGT

50

55

5

10

15

20

25

30

35

40

45

50

55

ACGCCCACCCGGCGGCCGTCTTCGTCTTGACGCCGGACGGGCGGATCTCC  
CGGTACCTGTACGGGACGGAGTTCCCGGCGCGCATCTCCGGCTCGCGCT  
CCTGGAGGCGAGCCGCGGCGGTATCCGCACGATCGTCGATCGGGTGATCA  
TGACCTGCTATCGCTTCGACCCGGCGAGCCGGAGATACGCTCCGTTCCCTA  
CTCGGCTTCCTCCGGCTCGGGGCGGCGGCCATCCTGATCACGGTCGGCGG  
GCTGCTCGCCGTCTGTGGCGGCGGAGCGCCGGCGGCCAGGTGCTCGCA  
CGAGCGCCCGCGTCGGTCGTGACGCCGTGGCCGACCGCCAGGGGAGGTCA  
CCATGATCAACGAGCTCCTGCGCAAGCTTCTTTTCTGTCCGGCCAGTGG  
TCGACGATCGTGTTTCGACATTTACAAGCTGCTTTACTTCGTGATCTCGGT  
GACGATGGCCGGCGCGACGCTCGTCGCCCTGTTTCGCGGCCTACCTGATGA  
TCCGGTACCGCAGGCGCCAGCGGGATGTTGAAGGCCCGTTCCCCGGAGCG  
ACCGCGAGGCCTCCGCTCCTCCTCGAGGTCCGCATGGTGCTGGGCCTCAT  
CGTCCTGTTCTTCGTCTGGTGGGTGATTTGGAATGCGGCAGTATGCAGAGC  
TCCGCGTCGCCCCCGCGGACCCGGTCGTGGTGTACGTGACCGGGAAGCAG  
TGGATGTGGAAGTTTCGCCTACCCGGAGGGCCCGAGCTCGGTGGCGACGCT  
CTATGTGCCGGCGCGTCGGCCGGTGAAGCTCGTCATGACGTCCCGGGACG  
TGATCCACAGCTTCTTCGTCCCCGATTTTCGCATCAAGTACGATGTTCGT  
CCCGGCCGCTACACCACGCTGTGGTTCGAGGCGACCGCGCCGGGCGCCTA  
TCAGATCCTGTGCACCGAGTACTGCGGGACGAACCACTCCACCATGCGCG  
GCGAGGTGATCGCGCTCGAGCCCTCCGATTTTCGCGCGGTGGCTCTCCGAC  
CGCGGGCGGGGCGCCGGTATCGCCGGACAGGAGTACACGCCGCCGTTCGAC  
GCCGGGCGAGGGGATCCCGCGCGAGCCGCTCAGCCTCGTCCGGCTGGGCG  
AGAACATCGCGGCCGAGGAGGGCTGCCTGCGCTGCCACACGCCGGACGGG  
ACACCGCACATCGGGCCGACCTGGGCCGGCCTCTACATGTCGGTCGTCCC  
GCTGGAGAGCGGCGGCGCCGGTTCGCCGACGACGCGTACATCACCGAGT  
CGATGATGGATCCGCTCGCCCGGATCCACCGCGGCTACCAGCGGGTCATG  
CCCTCGTTCCTCGGCCGGCTCCAGCCGGCGCAGGTCCCGCCATCGTCGA  
GTACATCCGGTCGTTGAGGGGCGTCGCGCCGGAGCCGGGCGCGCGGACGC  
CGCTGCCCCGAGGGCCCCGCCCTTCCTGCGCTCCGGCCCCGAGCGCCCCGCC  
CCGCTCAGCGGGGGCGCGCCGGTCCGCCCGATCGAGGGCGGCAAGCCCCGG



5

10

15

20

25

30

35

40

45

50

55

GGAGGAGCTCCGATGAGCACGGAAAGCGTACGAATCTCTGCCCCGACGCGCC  
GGCCGAGAGGCCCCGAGCCCCGACTACCTCCATGTTTACCGCGGGGTGACG  
GAGTGGCTCACGACCACGGATCACAGCGGATAGGTCTCATGTTCTACGC  
CGTCATCGTCGGGAAAGCTTCTTCTCGGAGGCATATTCGCCCTCATCAT  
5 GCGGACCGAGCTCCTCACGCCCCGAGCGGACCATCATCGACGCGGCGACCT  
ACAACCGGATGTTACGCTGCACGCGGTGATCATGGTCTGGCTGTTTCATG  
ATCCCGTCGATCCCCAACGCGTTCCGCAACTTCGTCTCGCGATCATGCT  
CGGCGCCAAGGACCTCGCGTTCCCCCGGATCAACCTCGCGAGCTTCTACA  
TCTACCTCCTCGGGGCGGCGATCGCGATGGGCGGCATGATCGCGGGCGGC  
10 ACGGACACCGGCTGGACGTTCTACCGGACGTACAGCCTGAAGACGCCGAT  
GACGCTGTTCCCGGTGCTCTTCGGCATCTTCATCGTCGGCGTCTCGTCCA  
TCATGACGGCGGTCAACTTCATCGTTGACCACGCACACGATGCGCGCCGAG  
GGGCTCACGTGGAGCCGCTGCCGCTCTTCGTCTGGAGCACCTACGCGAC  
15 GAGCATCATCTGCTCTTCGCGACCCCGGTCTCGGGCTCTCGATCCTGC  
TCATCGGCATCGACCACGTGACCGCGCTCGGGATGTTTCGATCCCCGGTTC  
GGCGGCGATCCGGTCTCTTCCAGCACCTCTTCTGGTTCTACTCCACCC  
CGCCGTCTACATCATGATCCTGCCCGCGTTCCGCGTGGTGAGCGAGGTCG  
20 TCTGCACGTTTCGCGCACAAGCGCCCCGCGTCTACTGGGCGATCGCCATC  
TCGTGCTCGGGATCGCGTTCGTGGGTTCTGGACGTGGGGCCACCACAT  
GTTCGTGGCGGGGATGAGCGAGTACCGCGCGGACGTCTTCGGCGTGCTCT  
35 CGATGTTTCGTGGCCATCTTCTCGGCCATCAAGGTCTACACGTGGGTCGCG  
ACGCTGTACAGGGGCTCGATCCACTTCAACACGCGCTGCTCTACTTCAT  
CGCCTTCCTCTTCTGTTCTGTTTCGGGGGATGACGGGCGTGCCGTCG  
CCACGCAGTCGCTGGACGTGCACTGGCACGACACATACTTCGTTGTGGCG  
40 CACTTCCACTTCATCATGGTGGGCGGGACGCTCACCATGTTCTCGCGGC  
GGCGCACTACTGGTTTCCGAAGATTTTCGGGCGCCTCTACTCGGAGCGCG  
TCGGGCTCCTCTCGGCCGCGTCCGTGTTCTCCTCGGCTTCTTCTTGACCTTC  
45 TTCCCGCAGTTCCTCCTCGGGAACATGGGGATGCCCCGCGCTATTACAG  
CTACCCGCGCGCTACCAGTGGCTCCACGTGCTCTCGACCGGCGGCGCCT  
30 ACCTGCTCGCCGCGGCGCTCGTGATCTCGCTCCTGAACCTCGTCATCGCG

5

10

15

20

25

30

35

40

45

50

55

CTCAAGTGGGGCCGGAAGGCCGGGAGGAACCCCTGGGGCGGGCGCACGCT  
CGAGTGGATGACCGGCGAGCCCTTGCCGCCCAAGCACAACTTCCCGGTTCG  
CGCCGCTCGTCCGCGCGGCCCGTACGAGTTCCAGCTCTCCGAGGAGGAC  
GCCCCGTGCGACAACCACGCCCGCTGCGTGAGCAGTTCGAAGATCTCGAGA  
5 AGCAGACGCACGCGGCCCGCCTCGGGATGTGGTTGTTTCCTCGGGAGCGAG  
GTGCTCCTCTTACCGGGCTCTTCGCGCTGTACGCGGCGTACCGCGAGCT  
15 CTACCCGCGCGATTTCGCCGAGGCCATCGCGCACAAACGTCGCGATCG  
GCACCACCATGACGCTCATCCTGATCGGCAGCAGCTTCACCGTCGCCATG  
GCGGTGCACGCCGTCCGCGCCTCCCACCCGCGGCGCGCCGCGCTGTTCTT  
20 10 CGCGGTGAGCGTGGCGATCGGGATCGTGTTCTCGTGCTGAAAGGGATCG  
AGTACGCGCAGCACTTCCGCGAGGCGCATCTTCCCGGCCGGCGCCTACCGC  
TTCGCGGAGCTCCCGACGTTCCGCGCGCAGATGGCGTTCACGCTGTACTT  
CGCCATGACGGCGCTCCACGCCCTGCACGTCGTGGGCGGGGCCGGCCTCC  
25 TCACGGGGGTGCGGTGGGGGTGCTGGAAGGGCCGGTACTGGGCATACGAC  
15 CAGACGCCCGGTGGAGCTCAGCGGCCTCTACTGGCACCTCGTGGACATCAT  
GTGGATCTTCATCTGGCCGCTCCTCTACCTGACGCGCAAGTGACGGCACA  
GCCCCGAGACGACCATGCCGCAAGAGCACGTGCGGAAAGCACGCCCTGG  
30 ACCCGTTACCTCATGGACGCTGATGGCCCTCATTGCCCTCACGCTCCTGT  
CGTTAGCGCTCTCGTTCTTGCGCACGGGGGCTTGGGAAATACCGATCGCG  
20 CTGCTCCTCCCCGTGGTGAAGAGCGTGCTCGGGCTTGCTTTCTTCTTGCT  
35 CCGTCGAGGTGCATTAAGGTCATCAAACGCCTTTTCAAACATTGGCTGC  
TGGTGTACTTACTTGGACGTTTCATACTGTCTCTTTTATGGCCCCCAACG  
GTCATTACGTCGACCTACACAATTTCTCTTCCCGCACCCCTCCTACCTGTA  
TCTCTAAGCACTGCCTTGCGTCCTGCTCTATTACATTCTACTCCGGCTGT  
40 25 CCATGTGTGGGATTATATGCGCGAGGTACCTATTCCGCCGTGGAGTCTCC  
ATTTACCTCTTGGACCTTGCCCGTTCTGAATCTCGATCTCCTCATGCGTT  
GGTCCACCATGTATTACCTCCTAGAATCTTATACTCCATATCTCTATATA  
45 TCTAGTTGTGCGTGTAATTGTGTCATATATTATCGCCACTGCTGTATGAA  
TACCGTGCCGACGTGCTATATACGAAANTACTCCTCGGTGATATCTCCA  
30 CCTCATATATACCTCCGAGTGTAGTATACGCACGAGTGTATATACTCTTC

5

10

CTCTGGTCACGCGACTTCGTGCTGATATGATACCATCGTTCCATGTTACG  
CGAAGTTACTCATAAGATCTCCTCACACATCAACGAGTGTACTCCTATGT  
GTTTCATACAACTCGATACCCTTCAGAGTAGTGTGCCTATGTGGTA  
TGCATAATGTTAGTATACTTT

5

15

SEQ ID No 79 (&gt;Contig54)

20

10

TGGGAAAGAGGGCCACAGGGGATGTAGCAGGACGCTTAATAGTAAATGAC  
GAGGGTGTGCCGACGAGACCCGTAGGAAACAACGGGCACAGACGAGAGCA  
ATAAAGGGGGTTTGAAGGTACCCCGGATAGAGTAGAGAAGGCTAGCGGAC  
GAGTAAGACGCGGAGGAAATAAGTCGGCGTCGTAGAAGTTCTGTGGAGAA  
GGTACGACTCTTAAAGACCTAGGCGGGAGACAGTTTCCACCCGAGGCAGA  
GCAAGACCACAAGATTCAGAGGGAGTAAGGAGTTCCGAATTGGAGAGGTT  
GAGGGGCGTGTGAGCCGTCAAGTGGGGCGCGTACGCAAAGAAAGAAGCGT  
CCATGTCAGAGGCCACGCGGCCGTTGCGGCCCTACATGGGTGATCACGGT  
GGGCCGTAGGCGGACCGGAGATGAGCGCGGCTCCGCCACCGGACGGCGA  
GGGGCACGGCGCCTTCGTCCGGGGCGCTCGCGCGCGATCTCGCGCGCGCG  
GGGGGCGCCGGCGTCCCGTCTCGGCGAEGGCCAGCCGCCCATCGTCCA  
CGCCCTCGGGCACGTCAACGCGCGCTCCGCAGCCGGCGGCCTGGA  
TGGTCGATCCTGTGCTGATCGACGCGGGCCCTCCACGCAGAGCTTCTCC  
GAGCTCGTCGGCGAGCTCGGGCGCGGCGCGGTGACACCTTGATCCTCCT  
CGACGTGAACCCCGTGTACGCCGCGCCGGCCGACGTGATTTGCGGGGCC  
TCCTCGCGCGCGTGCCACGAGCTTGAAGGCCGGGCTCTACGACGACGAG  
ACCGCCCGCGCTTGACGTGGTTGCTGCCGACCCGGCATTACCTCGAGTC  
GTGGGGGGACGCGCGGGCGTACGACGGGACGGTCTCGTTCGTGCAACCCC  
TCGTCCGGCCGCTGTTGACGGCCGGGCGGTGCCCGAGCTGCTCGCCGTC  
TTCGCGGGGGACGAGCGCCCGGATCCCCGGCTGCTGCTGCGCGAGCACTG  
GCGCGGCGAGCGCGGAGGGGCGGATTTGAGGCCTTCTGGGGCGAGGCAT  
TGAAGCGCGGCTTCCTCCCTGACAGCGCCCGGCGAGGCAGACACCGGAG  
CTCGCGCCGGCCGATCTCGCTCAGGAGCTCGCGCGGCTCGCCGCCGCGCC  
GCGGCCGGCCGGCGGCGCGCTCGACGTGGCGTTCTCAGGTCGCCGTCGC

50

55

5

10

15

20

25

30

35

40

45

50

55

TCCACGACGGCAGGTTTCGCCAACAAACCCCTGGCTGCAAGAGCTCCCGCGG  
CCGATCACCAGGCTCACCTGGGGCAACGCCGCCATGATGAGCGCGGCGAC  
CGCGGCGCGGCTCGGCGTCGAGCGCGGCGATGTCGTCGAGCTCGCGCTGC  
GCGGCCGCACGATCGAGATCCCGGCCGTCGTCGTCCGCGGGCACGCCGAC  
5 GACGTGATCAGCGTCGACCTCGGCTATGGGCGCGACGCCGGCGAGGAGGT  
CGCGCGCGGGGTGGGCGTGTGGGCGTATCGGATCCGCCCGTCCGACGCGC  
GGTGGTTCGCGGGGGGCCTCTCCGTGAGGAAGACCGGCGCCACGGCCGCG  
CTCGCGCAGGCCAGCTCGAGCTCTCCAGCACGACCGTCCCATCGCGCT  
CAGGAGGACGCTGCCGCAGTACCGTGAACAGCCCGGTTTCGCGGAGGAGC  
10 ACAAGGGGCCGCTCCGCTCGATCCTGCCGGAGGTCCAGCACACCGGCGCG  
CAATGGGCGATGTCCATCGACATGTCGATCTGCACCGGGTGCTCCTCGTG  
CGTCGTGGCCTGT CAGGCCGAGAACAACGTCTCTCGTCGTGCGCAAGGAGG  
AGGTGATGCACGGCCGCGAGATGCAGTGGTTGCGGATCGATCAGTACTTC  
GAGGGGGGAGGCGACGAGGTGAGCGTCGTCAACCAGCCGATGCTCTGCCA  
15 GCACTGCGAGAAGGCGCCGTGCGAGTACGTCTGTCCGGTGAACGCGACGG  
TCCACAGCCCCGACGGCCTCAACGAGATGATCTACAACCGATGCATCGGG  
ACGCGCTTTTGCTCCAACAACGTCCCGTACAAGATCCGGCGGTTCAATTT  
CTTCGACTACAATGCCACGTCCCGTACAACGCCGGCCTCCGCAAGCTCC  
AGCGCAACCCGGACGTGACCGTCCGCGCCCCGCGGCGTCATGGAGAAATGC  
20 ACGTACTGCGTGACGCGGATCCGAGAGGCGGACATCCGCGCGCAGATCGA  
GCGGCGGCCGCTCCGGCCGGGCGAGGTGGTCACCGCCTGCCAGCAGGCCT  
GTCCGACCGGCGCGATCCAGTTCCGGTTCGCTGGATCACGCGGATACCAAG  
ATGGTCGCGTGGCGCAGGGAGCCGCGCGGTACGCCGTGCTCCACGACCT  
CGGCACCCGGCCGCGGACGGAGTACCTCGCCAAGATCGAGAACCCGAACC  
25 CCGAGATTGAATGAGCCATGGCGGGCCCGCTCATCCTGGACGCACCGACC  
GACGATCAGCTGTGAAGCAGCTCCTCGAGCCGGTATGGAAGCCGCGCTC  
CCGGCTCGGCTGGATGCTCGCGTTCGGGCTCGCGCTCGGCGGCACGGGCC  
TGCTCTTCCTCGCGATCACCTACACCGTCTCACCGGGATCGGCGTGTGG  
GGCAACAACATCCCGGTGCGCTGGGCCTTCGCGATCACCAACTTCGTCTG  
30 GTGGATCGGGATCGGCCACGCCGGGACGTTTCATCTCCGCGATCCTCCTCC

5

10

15

20

25

30

35

40

45

50

55

TGCTCGAGCAGAAGTGGCGGACGAGCATCAACCGCTTCGCCGAGGCGATG  
ACGCTCTTCGCGGTCTGCCAGGCCGSCCTCTTTCGGTCTCCACCTCGG  
CCGCCCCCTGGTTTCGCCTACTGGATCTTCCCGTACCCCGCGACGATGCAGG  
TGTGGCCCGCAGTTCCGGAGCGCGCTGCCGTGGGACGCCGCCGCGATCGCG  
5 ACCTACTTCACGGTGTCTGCTCCTGTTCTGGTACATGGGCCTCGTCCCGGA  
TCTGGCGGCGCTGCGCGATCACGCCCGGGCCGCGTCCGGCGGGTGATCT  
15 ACGGGCTCATGTCTCGGCTGGCAGCGCGCCGCCGACCACTTCCGGCAT  
TACCGGGTGCTGTACGGGCTGCTCGCGGGGCTCGCGACGCCCTCGTCTGT  
CTCGGTGCACTCGATCGTGAGCAGCGATTTTCGCGATCGCCCTGGTCCCCG  
20 GCTGGCACTCGACGCTCTTTCGCGCGTTCTTCGTTCGCGGGCGCGATCTTC  
10 TCCGGGTTCGCGATGGTGCTCACGCTGCTCATCCCGGTGCGGCGGATCTA  
CGGGCTCCATAACGTCTGTGACCGCGCGCCACCTCGACGATCTCGCGAAGA  
TGACGCTCGTGACCGGCTGGATCGTCATCCTCTCGTACATCATCGAGAAC  
25 TTCCTCGCCTGGTACAGCGGCTCGGCGTACGAAATGCATCAGTTTTTTCA  
15 GACACGCCTGCGCGGCCCCGAACAACGCCGCTACTGGGCCCAGCACGTCT  
GCAACGTGCTCGTCATCCAGCTCCTCTGGAGCGAGCGGATCCGGACGAGC  
30 CCGTTCGCGCTCTGGCTCATCTCCATCCTCGTCAACGTCTGGGATGTGGAG  
CGAGCGGTTACGCTCATCGTGATGTCTGCTCGAGGAAGAGTTCCTCCCGT  
CCAAGTGGCACGGCTACAGCCCGACCTGGGTGGACTGGAGCCTCTTCATC  
20 GGGTCAGGCGGCTTCTTCATGCTCCTGTTCTGAGCTTTTTGCGCGTCTT  
35 TCCGTTTCATCCCCGTGCGGAGGTTCAAGGAGCTCAACCATGAAGAGCTGG  
AGAAGGCTCGGGGCAAGGGGGGGCCTGATGGAGACCGGAACGCTCGGCG  
AGTTCGACGACCCGGAGGCGATGCTCCATGCGATCCGAGAGCTCAGGCGG  
CGCGGCTACCGCCGGGTGGAAGCGTTACGCCCTATCCGGTGAAGGGGCT  
40 CGACGAGGCGCTCGACCTCCCGCCTCGAACCTCAACCGGATGGTGCTGC  
25 CCTTCGCGATCCTGGGGGTCTGTGGGCGGCTACTTCGTCCAGTGGTTCTGC  
AACGCTTTCACCTATCCGCTGAACGTGGGCGGGCGCCCGCTGAACTCGGC  
GCCGGCGTTTCATCCCGATCACGTTTCGAGATGGGGGTGCTCTCCACCTCGA  
45 TCTTCGGCGTGCTCATCGGCTTTTACCTGACGAGGCTGCCGAGGCTCTAC  
30 CTCCCGCTCTTCGACGCCCCGGGCTTCGAGCGCGTCACGCTGGATCGGGT

5

10

15

20

25

30

35

40

45

50

55

CCTGGTCGGGCTCGACGACACGGAACCTTCCTTCTCGAGCGCCCAGGCGG  
AGCGCGATCTCCTCGCGCTCGGCGCCAGGCGCGTCTCGTGGCGAGGAGG  
CGCGAGGAGCCATGAGGGCCGGCGCCCCGGCTCGCCCCCTCGGGCGCGCG  
CTCGCGCCGTTCGCCCTCGTCCTGCTCGCCGGGTGCCGCGAGAAGGTGCT  
5 GCGGAGCCGGACTTCGAGCGGATGATCCGCCAGGAGAAATACGGGCTCT  
GGGAGCCGTGCGAGCACTTCGACGACGGCCGCGGATGCAGCACCCGCCC  
GAGGGGACCGTCGCGCGCGGGCGCGTCAACGGGCGCCCGGCTATCTCCA  
GGGCGTCCTCGACGGGGCGTACGTACGGAGGTGCCGCTCTCGCTCACGG  
TCGAGCTCGTGCAGCGCGGCCGGCAGCGCTTCGAGACCTTCTGCGCGCCG  
10 TGCCACGGGATCCTCGGCGACGGCAGCTCGCGCGTGGCGACGAACATGAC  
GCTGCGCCCGCCGCGTCTCGTCTCGGACCCGAGGCGCGGAGCTTCCCGC  
CGGGCAGGATCTACCAGGTCATCATCGAGGGCTACGGCCTGATGCCGCGC  
TACTCGGACGATCTGCCCAGACATCGAAGAGCGCTGGGCCGTCTGCGCCTA  
25 CGTGAAGGCGCTTCAGCTGAGCCGCGGAGTGGCCGCGGGCGCCCTCCCGC  
15 CCGCGCTCCGCGGCCGGGCAGAGCAGGAGCTGCGATGAACAGGGATGCCA  
TCTAGTACAAGGGCGGCGCGACGATCGCGGCCTCGCTCGCGATCGCGGCG  
CTCGGCGCGGTTCGCCGCGATCGTCTGGCGGCTTCGTCTGATCTCCGCCGGTT  
CTTCTTCTCGTACCTCGCCGCGTGGTCTGTTCCGCGTATTCTGTCCGTGG  
GCGCGCTCGTCACGCTCCTCACCTGCAACGCCATGCGCGCGGGCTGGCCC  
20 ACGGCGGTGCGCCGCTCCTCGAGACGATGGTGGCGCCGCTGCCCCTGCT  
35 CGCGGCGCTCTCCGCGCCGATCCTGGTCTGGCCTGGACACGCTGTACCCGT  
GGATGCACCCCGAGCGGATCGCCGGCGAGCACGCGCGGCGCATCCTCGAG  
CACAGAACGCCCTACTTCAATCCAGGCTTCTTCGTCTGCGCTCGGCGAT  
CTACTTCGCGATCTGGATCGCCGTCGCCCTCGTGCTCCGCCGGCGATCGT  
40 TCGCGCAGGACCGTGAGCCGAGGGCCGACGTCAAGGACGCGATGTATGGC  
25 CTGAGCGGCGCCATGCTGCCGGTCTGTGGCGATCACGATCGTCTTCTCGTC  
GTTCTGACTGGCTCATGTCCCTCGACGCGACCTGGTACTCGACGATGTTCC  
CGGTCTACGTGTTCTCGGAGCGCCTTCGTGACCGCCGTCGGCGCGCTCACG  
45 GTCCCTCTCGTATGCCGCGCAGACGTCCGGTTACCTCGCGAGGCTCAACGA  
30 CTCGCACTATTACGCGCTCGGGCGGCTGCTCCTCGCGTTACGATATTCT

5

10

15

20

25

30

35

40

45

50

55

GGGCCTATGCGGCCTATTTCCAGTTCATGTTGATCTGGATCGCGAACAAG  
CCCGACGAGGTCGCCTTCTTCCTCGACCGCTGGGAAGGGCCCTGGCGGCC  
GACCTCCGTGCTCGTCGTCCCTACGCGGTTTCGTCTCCCGTTCCTGATCC  
TGATGTCGTACGCGATCAAGCGGCGCCCGCGCCAGCTCTCGTGGATGGCG  
CTCTGGGTCGTGCGCTCCGGCTACATCGACTTTCACTGGCTCGTGGTGCC  
GGCGACAGGGCGCCACGGTTTCGCTATCACTGGCTCGACCTCGCGACCC  
TGTGCGTCGTGGGCGGCTCTCGACCGCGTTCGCCGCGTGGCGGCTGCGA  
GGGCGGCCCGGTGGTCCCGGTCCACGACCCGCGGCTCGAAGAGGCCTTTGC  
GTACCGGAGCATATGATGTTCCGTTTCCGTACAGCGAGGTTGCCAGGA  
GGAGGACACGCTCCCCTGGGGGCGCGTGATCCTCGCGTTCCCGTCTGTGC  
TCGCGATCGGCGGCGCGCTGACGCTCTGGGCCTGGCTCGCGATGCGGGCC  
CGCGAGGCGGATCTGCGGCCCTCCCTCGCGTTCCCCGAGAAGGATCTCGG  
GCCGCGGCGCGAGGTCAGCATGGTCCAGCAGTCGCTGTTTCGACGAGGCGC  
GCCTGGGCCAGCAGCTCGTCGAGGCGCAGCGCGCGGAGCTCCGCCGCTTC  
GGCGTCGTGATCGGGAGAGGGGCATCGTGAGCATCCCGATCGACGACGC  
GATCGAGCTCATGGTGGCGGAGGGCGCGGATGAGCCGGGCCGTGCCGT  
GGCCCTCCTGCTGGCAGCCGGCCTCGTGTCGCGCCCGGGCGCCGCTCCG  
AGCCCGTATCGCTTTCGCCCCGCGTGGGCCCGTCCGCGGGCGAGGCCGC  
GCTCTGAAACGACGGCTCCGGCGCGGATGAGCGGCCCGAGGCGACCTCCT  
GCAACCCACCGCGCTGCGTACAGGGGTAAATCAACTGCATTCCAGGATAC  
GGACCGCGCAGCATAAACCTCACCAGGACAGACTGAATAATGCCGAACCT  
TGAACCTCTATGCCCATGCGTGCGGGATCATCATGCCCAACTATTTAAAC  
TGGTCCCTCCGGCAAAAGGAACGGACCAGCACCCAGAATAACCCCTGTTT  
GGCCAGCGCAAAAAAATGAAAACTCTTCTCCTTGCGGCTAAAATAAACGA  
CTCCGGGGAACCGAATGGTATGCAATACACCATGGCAACGCGATTCTGTCC  
TAACCTTAAACAACTTCCCAGCGACTCGTCCGTGCAACGCTTCTGACGC  
AAGACCATGACCCACGAGAACCGGGCGGCGGACACACTGCCAGTGAAAC  
TCGGCCTAGGCCCGCCCTGCCTTCACGTATTCACCGTGGGGCGGTCCAAA  
ACTAAAACAATACGTGACTCTTCTCAAATATCGTCGGATAAAGGCCAAC  
ACGCGTACCTCCCCCTAAAGGGAAGAAACCCCTACCAGGGTGGACCGTAT

5

CCACCCGTGATCCCTGAAACATCTCCAGCCGCGTACAAATTAGGCTTTGA  
CAAAACC

10

SEQ ID No 80 (>Contig55)

5 GGGGAAAGAAGTCGGGAGCAGCAGAAGAGCGCAGCGATTGAAAGCACGAC  
15 GCAAGGGCAAAGGACATNAGACGACAACAGAAAGGCGAAGACATGAGAAG  
AGGAGAAAGACTCGGAAATGGGCACAAAGCGAGAAAAACGTAACGCTAT  
CTGAGAAAGAACACCATAAAGTCAGACTGGGGTAAAGCCATACACGCAGC  
GAAAGCAGACAAGCAAAAGTCATTAACAGATGCAACAAAGAACAAAAGAG  
20 10 AAAGGAAGGACACATGAAGGAGAAAGGCGGCTTCGAGAACTAAAGACGGG  
AGCCAACAATATACCTTTTCACAGGGGCGAAGAAGGGCCCAAGGTTCAATC  
GAAGACGATTGAATCCAAGGAAGTCCAATCGGAAGAAAAAGCATATTGA  
TAAACCAGACAACGACAGCAGGCCTGAACGTAGGAGCGAGATCGTGAGAC  
25 ATCAGTAGGCAAAACAAGAGCGCTACACCCAGGGGTCGTCAACCTAGAAA  
15 GCGCGTCCTCAAGCCGGTAGCGGCCGCGCGCGACCCAGCCCGATGCGGGC  
GCCTTCGCGCGCGAACC GCCGCGACCGTCGCGCGCCCCACGCCCGCTGACG  
CCCCCGTGATCACCACGACCTCCGGCCGCTGGAATCGCCCATGCCCCGTC  
30 GCCTCCCGCCTCGCCCCGCGCGTCAAGCAACGTGAATGCCACCTGAGCGT  
GTCACTTCCTCAAGCTCGACAGCACGTCCTTGATCCGCTCGGTCCGACCC  
20 GCCGTGAGCGACCCAGGGGCTCGGCTGCCGCACGACGTGCAGGTCCGCTCG  
35 ACGCTCGAGCACCTCGAACCGCGTCGTTCCCTGCTCGGTCCGGACGAACC  
GGAGCGCCACCACGGCCGTGCCACATGGAGGTTGAGAGGGTGATCTCG  
GGCAGCCACGCCGGCAGCGCCGGGTCGACCAGCAGCAGGTGAGCGGCGC  
GAACGGGTACAGCCCGAGCATCGCCTGGAGCATCGCGAACACGCTCGAGC  
40 25 AGGACCAGGCCTGCGGCCAGTTGGCCTTCGGGTACAGCGCGGGGAACGGG  
TGCTCCGCGTCGCGCGGGTGGCCGCTCCAGCACTCGGGCAGCCGGTGGTG  
CTCGAACAGCGCCGCCGCTCGAACACGGCGCGGCACAGGAGCGCCACGT  
45 GCCCGTGACAGGCCGTACCGCGCGAGCCCGAGCGCGATCGCGCCCTGATCG  
ACGGGCCAGACCGTCCCCCGGTGATAGCTGTACGGATCGAACCGGGGTG  
30 ATCGGCCGAGAGCGTGCGGATACCCCGAGCCGGAGAACATGTCCGCCGCGA

50

55



160

5

10

15

20

25

30

35

40

45

50

55

ACATGCGGATGGCGGTGGGCTCGGCGAGGGCCTGGTCCACGATCCCCGCC  
GCGAGGCAGAGCCCCGGATCCGAGCCGATCGAGCGGATCTGGCGCTTGTC  
CGGGCCGAGGCCCATCGCGAAGGTGCGCGCGTCGGGCATCCAGAAGGCGT  
CGTTGAACCGCCTCTGCAGCTCGAGCGCCTCGGCGAACAGCCGTCGCGCG  
5 TCGTCCTTGCGGCCGAACCAGAAGAGCAGCTCGGAGAGGCGCAGCTTCGA  
CAGGAACACGAAGCCCTGCATCTCGCACGTCCCGATCGGCGGCCGACCT  
GAGAGCCGTCGGCGTGGACGATGGCGTCGTCGGAGTCCTTCCAGCCCTGG  
TTCTGGATCGACGCGCTCGAGCGGGGCTCGTACTCGTAGAACCCGTCGCC  
GTCGAGATCGCCCTCCTCGTCGATCCAGCGCATGGCCCTGAGCGCAGGCT  
10 CGATCAGGCGGCCGACGCGCTCGCGATCGCCGGTCCAGTGCCAGAGCTCC  
GAGGCGGCCACCCGCTAGAACATCGTCGAGGTGCGCGACGCGTACGTGCG  
CCCCAGCGGGTTGTAGTTGAGGTCCGAGAGCGCTCCGTCCCTGGCCTGAT  
GCAGCATCCGGTCGGGCTGCTCGTCGCGCCAGTCGTCGACGACGCGCCCC  
TGCCAGCGCGGGAGCACGAGCGCTGTGCCTGCGAGGATGTCGGTCGTCAG  
15 CGCGGCGGCCTGCGTGCCCGCGGCGAGCGGGTCGCGGCCGAAGAGCCCGA  
TGTAGATCGGCAGTCCGGCGGCCACCGTCCAGGAGCGCTCGTCCTGGTCG  
ATATCGTACATGCGGAGCGCGATGAGATCGCGCTTGGCTCGCTCGAGCAC  
GGAGAAGACCGTATGCGAGAGCGTGTGCGCGCCCGGGACCGAGAAGGACG  
TGGCGCGGTGGTGAAGGTGCCGCGCGCCACGTCGCGCGCGTTTCGCGTG  
20 CCGAAGAACGAGCGGCAGCCGGCGAGCAGCGGGAGCGGTTCGCCGCGGAT  
CAGGGCGATCACGTCGACGCAGCCGCGCCAGGTGCCATGGGGCTCGAGAT  
CGATCGAGAACCGGATCTCGCGGCCAGCGCAGGACGGGGGCGAGCCCGCG  
CTCCGGGCCCCGTACGACGATCCCGGCGTCGAAGCGCGCGACGCCGGATT  
TCCGGGGTGTTATAGCGGTGCTCGGCGCGGTAGTCGCAGCGCAGCGCGC  
25 ATCCGTGCTCGGCGGGCTCGAGCGCCACCGGGCGTCGCCGCGCTGGAGG  
CGCGGGCCGTCCGTCTCCTCCGCGTCGGCGAAGTCGGCGTCGATCTCGAG  
CGCCAGCGTGAAGCGCACGCGCTCCTGCGTGAAGCTCGCGACGTCGATGT  
CCTCGTGGAAGCCGTCCCCACGAAGCGCGACAGCCGAGCTCGACGGTC  
CGCGGGCGCGCGCCCGCCCTCGGGCGGGCGGGGCGATGTAGTAGCCGAG  
30 CCAGCTGTCCGGCTCGACGGCGGAGAGCGCGACGGGCGCCGGCGGCCTGC

5

10

15

20

25

30

35

40

45

50

55

CGTCGATCAGGTGGCGGTAGAGGGAGAGGAGGCGGGTGTGCGCACGAAC  
AGGCCGATGTGCGGCTCGGGGGCGATCGAGCCGTCCGGACGCATGCACAG  
GACCGTCCGGTTCTGGCTGATGTACAGGGAGCCGGCGGGGGCCTCAGCG  
TGGCCAGCGAGCCGAAGGGTCTCTCGAGCGACATGGTCACCTCCAGGCGA  
5 GCCCGGGGTGAGGACTGCCAGGGGCGTGCCAACGACGTGGACGCGCTCAC  
GCCAAGCCGCTGGGCGGGCGCGGCGGCCCTGCGCCCGAGGCTCAGCCGA  
15 GCGCGAGCGGTGTTTCGTGCGGCGGAGCGGGCCTCGTCGTCGCCTCCAGG  
TAGGCCTTTCCGATGTGCGCCTCGTCCACGAACCCGACGATCGCGCCCTC  
GCCGTCGACGACGGGGACCTCGCGGACGCCGTGCGCCACCATCGCCTCGG  
20 TCGCCGTCCGCAGATCGTCGGTGACCGTCACGGCCACCGGCGGCTGCATC  
GCGTCGGCGGCCACGGTCATCCGCTCGAGGTCGTGCTCCACCGCGATGAT  
CCGGAGCGACTCGGCGGTGATCATGCCGACCATCTTGCGCGACGGTTCGA  
GCACCGGGAACACCTCCTGCCAGCTCGCGTCGGCCGCCCCGCGGAGCATC  
25 TCGCGGGCCGGCGTCCCCGGCACGAACGTACGAGCGCGCGCCCCTCGAT  
15 CATGATCTGCCGCACGCGGATGGTCTTGAGCACGTGAGCGTCGGCACCG  
GGTGCGCGGGAGACTCGCGCTGGGTGGGGAGCTGCGCGTGGTAGAGCGAG  
TGCTTCCGCAGCGCGACGAAGGCGACGCCCTCGGCGAGCATCAGCGGAAC  
30 CAGGAGGTCGTAGCTGCCGGCGAGCTCGCAGACCATCACGAGGGAGCTCA  
CCGGCACGTGCGCGACGCCGCCGTAGAAGGTGCCCATGCCACGAGCGCG  
20 AAGGCGCCCCGGGTGATGCGCGGATCGCCGAGCAGGAGCGCCGCCGCGCG  
35 CCCGAACGCGCCGCCGAAGAGCCCCGCCGATGACGAGCGACGGCGCGAAGT  
CGCCGGCGCACCCGCCGTGCCGAGCGTGAGCGACGAGGCGACGATCTTG  
GCGGCGCAGAGCAGGAGCAGGAGCTCCACGCCGCGCCAGCCCGGGTGGAG  
CCACGTGGCGCCGGTGATCGCGACCTGGACGGCGCCGTACCCGCCGCCGA  
40 GCAGCCCGAGCCCTTGCCCGGGGCTCTCGATCCTGCGGCCGACGAACCAG  
25 AGGACGGGCACGCAGAAGAGGCCAGCGCGAGCCCGCCGAGCCCCGGGCG  
CGCCCAGGGGGCGATGGGCAGGCGCGCCGCGATCGCCTTCACGCCGCCGA  
GGCACTTCAGGAAGCCGATCGCGACGAAGGCCAGGAGCAGCGCGAGCAGC  
45 GCATAGAGAGGGAGGTGCGACGGGACGAACGCATACTTCGGCGCGTGCGC  
30 GAAGAGCGTCGACTCGCCGTAGAACGAGATGAAGACCGAGTAGGAGACCA

5

10

15

20

25

CGCTGGCGAGCAGCGCCGGGATCAGCGCCTCGGCCTCGAAGTCGTCGCGG  
TAGAGCACCTCGACGGCGAGCAGGGCGGCGCCGAGCGGCGTGCGGAAGAT  
GGCGGACATCCCGGCCGCGACCCCCGCGAGCATCAGGATGCGGTGCTCGC  
GCCGGCCGACCGCGAGCCCGCGCCCCACGAGCGAGCCGAGCGCTCCTCCG  
5 ACCTGCATGGTCGGCCCCCTCGCGACCGCCGGCGCCCGCCGAGCCGAGCGT  
GAGGATCGACGCGACCGCCTTGACCCACGCGACCCGCTTGCGCATCCGGC  
CGCCGTGGTGGTGGAAAGGCCTGGATCATCGCGTCGCCGCCGCCGCCCGCG  
GCCTCGGGGGCGAGGCGCCAGGTGAGGATGCCCCCGGCCAGCGCTCCGAG  
CGCCGGGATCAGCAGCAGCAGCCAGAGCCGGACGCTCCGGTGCTCGGGGC  
10 CGTCGCCCGCTGAAGATGGCCTCGCCGTGGGCTCGAAGCCGCGCGTAGCCG  
GCGAGGCGGCCGAGGAGCAGCTCCTCGACGAGCTCGAGGGCGCCGAAGAA  
GAGCACCGCGACGAGGCCCGCGATGGCCCCGACGAGCACGGCATGCAGGA  
TCGTGCGCCCGACGAGCCTGAGATCGAGGGGCGCGACCTCCGAGAGGAGC  
25 GCGGAGAAGGGGCGCCGCCGTGCAACCACGCCGGCGCGCTCTGCTGGGAA  
15 TGGTTCCGTACGGCGATGGT

30

35

40

45

50

55

SEQ ID No 81 (>Contig56)  
GGATCCGGCCGCGAGCGGCTGCGGTGCCGATGCGCTCGACGGTGACGGGC  
GGGGTGATCGCGGGTCCGGAGCTCGGTGCGAGCTACTGGGCGGACAACCT  
20 TCGGCAGCCGGTGCGCTTCGCTGCGGCGGCGCAAGCGCTGCTGGAGGGTG  
GCCCCGCGCTGTTTCATCGAGATGAGCCCGCACCCGATCCTGGTGCCGCCC  
CTGGACGAGATCCAGACGGCGGGCCGAGCAAGGGGGCGCTGCGGTGGGCTC  
GCTGCGGCGAGGGCAGGACGAGCGCGCGACGCTGCTGGAGGCGCTGGGGA  
CGCTGTGGGCGTCCGGCTATCCGGTGAGCTGGGCTCGGCTGTTCCCCGCG  
40 GGCGGCAGGCGGGTTCCGCTGCCGACCTATCCCTGGCAGCACGAGCGGTG  
25 CTGGATCGAGGTCGAGCCTGACGCCCCGCCGCTCGCCGCAGCCGACCCCA  
CCAAGGACTGGTTCTACCGAACGGACTGGCCCCGAGGTGCCCCGCGCCGCC  
CCGAAATCGGAGACAGCTCATGGGAGCTGGCTGCTGTTGGCCGACAGGGG  
45 TGGGGTCGGTGAGGCGGTGCTGCGAGCGCTGTCGACGCGCGGACTTTCCT  
30 GCACCGTGCTTCATGCGTCGGCTGACGCCTCCACCGTCGCCGAGCAGGTA

5

10

15

20

25

30

35

40

45

50

55

TCCGAAGCTGCCAGTCGCCCAGAAACGACTGGCAGGGAGTCCTCTACCTGTG  
GGGCCTCGACGCCGTCGTCGATGCTGGGGCATCGGCCGACGAAGTCAGCG  
AGGCTACCCGCCGTGCCACCGCACCCGTCCTTGGGCTGGTTCGATTCTTG  
AGCGCTGCGCCCCATCCTCCTCGCTTCTGGGTGGTGACCCGCGGGGCATG  
5 CACGGTGGGCGGCGAGCCAGAGGCCTCTCTTTGCCAAGCGGCGTTGTGGG  
GCCTCGCGCGCGTCGCGGCGCTGGAGCACCCCGCTGCCTGGGGTGGCCTC  
GTGGACCTGGATCCTCAGAAGAGCCCGACGGAGATCGAGCCCCTGGTGGC  
CGAGCTGCTTTCGCCGGACGCCGAGGATCAACTGGCGTTCCGCAGCGGTC  
GCAGGCACGCAGCACGCCTTGTAGCCGCCCGCCGGAGGGCGACGTCGCA  
10 CCGATATCGCTGTCCGCGGAGGGGAGCTACCTGGTGACGGGCGGGCTGGG  
TGGCCTTGGTCTGCTCGTGGCTCGGTGGCTGGTGGAGCGGGGAGCTCGAC  
ATCTGGTGCTCACCAGCCGGCACGGGCTGCCAGAGCGACAGGCGTCGGGC  
GGAGAGCAGCCGCCGGAGGGCCGCGCGCATCGCAGCGGTGAGGGGCT  
GGAAGCGCAGGGCGCGCGGGTGACCGTGGCAGCGGTGGATGTCCCCGAGG  
15 CCGATCCCATGACGGCGCTGCTGGCCGCCATCGAGCCCCCGTTGCGCGGG  
GTGGTGACGCCGCCGGCGTCTTCCCGTGCGTCACTGGCGGAGACGGA  
CGAGGCCCTGCTGGAGTCGGTGTCTCCGTCCCAAGGTGGCCGGGAGCTGGC  
TGCTGCACCGGCTGCTGCGCGACCGGCCTCTCGACCTGTTCTGTGCTGTTT  
TCGTGCGGCGCGGCGGTGTGGGGTGGCAAAGGCCAAGGCGCATACGCCGC  
20 GGCCAATGCGTTCTTCGACGGGCTCGCGCACCATCGCCGCGCGCACTCGC  
TGCCGGCGTTGAGCCTCGCCTGGGGCTTATGGGCCGAGGGAGGCATGGTT  
GATGCAAAGGCTCATGCACGTCTGAGCGACATCGGGGTCTTGCCCATGGC  
CACGGGGCCGGCCTTGTCGGCGCTGGAGCGCCTGGTGAACACCAGCGCTG  
TCCAGCGTTCGGTCACACGGATGGACTGGGCGCGCTTCGCGCCGGTCTAT  
40 GCCGCGCGAGGGCGGCGCAACTTGCTTTCGGCTCTGGTCGCGGAGGACGA  
GCGCGCTGCGTCTCCCCCGGTGCCGACGGCAAACCGGATCTGGCGCGGCC  
TGTCGTTGCGGAGAGCCGCTCAGCCCTCTACGAGCTCGTTGCGCGCATC  
GTCGCCCCGGTGCTGGGCTTCTCCGACCCGGGCGCGCTCGACGTCGGCCG  
45 AGGCTTCGCCGAGCAGGGGCTCGACTCCCTGATGGCTCTGGAGATCCGTA  
30 ACCGCCTTCAGCGCGAGCTGGGCGAACGGCTGTGCGCGACTCTGGCCTTC

5

10

15

20

25

30

35

40

45

50

55

GACCACCCGACGGTGGAGCGGCTGGTGGCGCATCTCCTCACCGACGTGCT  
GAAGCTGGAGGACCGGAGCGACACCCGGCACATCCGGTCGGTGGCGGCGG  
ATGACGACATCGCCATCGTCGGTGCCGCCTGCCGGTTCACAGGTGGGGAT  
GAGGGCCTGGAGACATACTGGCGGCATCTGGCCGAGGGCATGGTGGTCAG  
CACCGAGGTGCCAGCCGACCGGTGCGCGCGGCGGACTGGTACGACCCCG  
ATCCGGAGGTTCCGGGCCGGACCTATGTGGCCAAGGGTGCCTTCCTCCGC  
GATGTGCGCAGCTTGGATGCGGCGTTCTTCGCCATTTCCCCTCGTGAGGC  
GATGAGCCTGGACCCGCAACAGCGCTGTTGCTGGAGGTGAGCTGGGAGG  
CGATCGAGCGCGCTGGCCAGGACCCGATGGCGCTGCGCGAGAGCGCCACG  
GGCGTGTTTCGTGGGCATGATCGGGAGCGAGCACGCCGAGCGGGTGCAGGG  
CCTCGACGACGACGCGGCGTTGCTGTACGGCACCAACCGGCAACCTGCTCA  
GCGTCGCGCTGGACGGCTGTGCTTCTTCCTGGGTCTGCACGGCCCGACG  
ATGACGGTGGACACCGCCTGCTCGTCGTCGCTGGTGGCGTTGCACCTCGC  
CTGCCAGAGCCTGCGATTGGGCGAGTGCAGACCAGGCCCTGGCCGGCGGGT  
CCAGCGTGCTTTTGTGCGCCGCGGTTCATTCGTGCGGGCGTCGCGCATGCGT  
TTGCTTTCCGACAGATGGGCGGTGCAAGACGTTCTCGGCCGCTGCAGACGG  
CTTTGCGCGGGCCGAGGGCTGCGCGGTGGTGGTCAAGCGGCTCCGTG  
ACGCGCAGCGCGACCGCGACCCCATCCTGGCGGTGGTCAAGAGCACGGCG  
ATCAACCACGATGGCCCCGAGCAGCGGGCTCACGGTGGCCAGCGGTCCTGC  
CCAGCAGGCGTTGCTACGCCAGGCGCTGGCGCAAGCGGGCGTGGCGCCGG  
CCGAGGTGATTTTCGTGGAGTGCCACGGGACGGGGACAGCGCTGGGTGAC  
CCGATCGAGGTGCAGGCGCTGGGCGCGGTGTACGGGCGGGGCGCCCCGC  
GGAGCGGCCGCTCTGGCTGGGCGCTGTCAAGGCCAACCTCGGCCACCTGG  
AGGCCGCGGGCGGCTTGGCCGGCGTGCTCAAGGTGCTCTTGGCGCTGGAG  
CACGAGCAGATTCCGGCTCAACCGAGCTCGACGAGCTCAACCCGCACAT  
CCCGTGGGCAGAGCTGCCAGTGGCCGTTGTCCGAGGGCGGTCCCCTGGC  
CGCGCGGCGCGCGCCCGCTCGTCCAGGCGTGAGCGCTTTCGGCCTGAGC  
GGGACCAACGCGCATGTGGTGTGGAGGAGCGCGGCGGTGGAGCCTGT  
GGCCGCGGCCCCCGAGCGCGCAGCGAGCTGTTCGTCTGTGGCGAAGA  
GCGCGGCGGCGCTGGATGCGCAGCTAGCCCGGCTGCGGGACCACCTGGAG

5

10

15

20

25

30

35

40

45

50

55

AAGCATGTCGAGCTTGGCCTCGGCGATGTGGCGTTCAGCCTGGCGACGAC  
GCGCAGCGCGATGGAGCACCGGCTGGCGGTGGCCGCGAGCTCGCGCGAGG  
CGCTGCGAGGGGCGCTTTCGGCCGCGAGCGCAGGGGCACACGCCGCCGGA  
GCCGTGCGTGGGCGGGCCTCGGGCGGCAGCGCGCCGAAGGTGGTCTTCGT  
5 GTTTCCCGGCCAGGGCTCGCAGTGGGTGGGCATGGGCCGAAAGCTCATGG  
CCGAAGAGCCGGTCTTCCGGGCGGCGCTGGAGGGTTGCGACCGGGCCATC  
15 GAGGCGGAAGCGGGCTGGTCTGCTGCTCGGGAGCTCTCCGCCGACGAGGC  
CGCCTCGCAGCTCGGGCGCATCGACGTGGTTCAGCCGGTGCTGTTCCGCA  
TGGAAGTAGCGCTTCTGCGCTGTGGCGGTCTGTTGGGAGTGGAGCCGGAA  
20 GCGGTGGTGGGCCACAGCATGGGCGAGGTGCGCGCGCGCACGTGGCCGG  
CGCGCTGTGCTCGAGGACGCGGTGGCGATCATCTGCCGGCGCAGCCGGC  
TGCTGCGGCGGATCAGCGGTGAGGGGAGATGGCGCTGGTTCGAGCTGTCTG  
25 CTGGAGGAGGCCGAGGCGGCGCTGCGTGGCCATGAGGGTCGGCTGAGCGT  
GGCGGTGAGCAACAGCCCGCGCTCGACCGTGCTCGCCGGCGAGCCGGCGG  
15 CGCTCTCGGAGGTGCTGGCGGCGCTGACGGCCAAGGGGGTGTCTGGCGG  
CAGGTGAAGGTGGACGTGCCAGCCATAGCCCGCAGGTGACCCGCTGCG  
CGAAGAGCTGATCGCGGCGCTGGGAGCGATCCGGCCGCGAGCGGCTGCGG  
30 TGCCGATGCGCTCGACGGTGACGGGCGGGGTGATCGCGGGTCCGGAGCTC  
GGTTCGAGCTACTGGGCGGACAACCTTCGGCAGCCGGTGCGCTTCGCTGC  
20 GCGGGCGCAAGCGCTGCTGGAGGGTGGCCCCGCGCTGTTTCATCGAGATGA  
GCCCCGACCCGATCCTGGTGCCGCCCTGGACGAGATCCAGACGGCGGGC  
35 GAGCAAGGGGGCGCTGCGGTGGGCTCGCTGCGGCGAGGGCAGGACGAGCG  
CGCGACGCTGCTGGAGGCGCTGGGGACGCTGTGGGCGTCCGGCTATCCGG  
TGAGCTGGGCTCGGCTGTTCCCCGCGGGCGGCAGGCGGGTTCCGCTGCCG  
40 ACCTATCCCTGGCAGCACGAGCGGTACTGGATCGAGGACAGCGTGCATGG  
25 GTCGAAGCCCTCGCTGCGGCTTCGGCAGCTTCGCAACGGCGCCACGGACC  
ATCCGCTGCTCGGGGCTCCATTGCTCGTCTCGGCGCGACCCGGAGCTCAC  
45 TTGTGGGAGCAAGCGCTGAGCGACGAGAGGCTATCCTACCTTTCGGAACA  
TAGGGTCCATGGCGAAGCCGTGTTGCCAGCGCGGCGTATGTAGAGATGG  
30 CGCTCGCCGCCGGCGTAGATCTCTATGGCACGGCGACGCTGGTGCTGGAG

5

10

15

20

25

30

35

40

45

50

55

CAGCTGGCGCTCGAGCGAGCCCTCGCCGTGCCCTCCGAAGGCGGACGCAT  
CGTGCAAGTGGCCCTCAGCGAAGAAGGTCCCGGTGGGCCTCATTCCAGG  
TATCGAGTCGTGAGGAGGCAGGTAGGAGCTGGGTGCGGCACGCCACGGGG  
CACGTGTGTAGCGGCCAGAGCTCAGCGGTGGGAGCGTTGAAGGAAGCTCC  
GTGGGAGATTCAACGGCGATGTCCGAGCGTCCTGTCGTCGGAGGCGCTCT  
ATCCGCTGCTCAACGAGCACGCCCTCGACTATGGTCCCTGCTTCCAGGGC  
GTGGAGCAGGTGTGGCTCGGCACGGGGGAGGTGCTCGGCCGGGTACGCTT  
GCCAGGAGACATGGCATCCTCAAGTGGCGCCTACCGGATTCATCCCGCCT  
TGTTGGATGCATGTTTTAGGTGCTGACAGCGCTGCTCACCACGCCGGAA  
TCCATCGAGATTCGGAGGCGGCTGACGGATCTCCACGAACCGGATCTCCC  
GCGGTCCAGGGCTCCGGTGAATCAAGCGGTGAGTGACACCTGGCTGTGGG  
ACGCCGCGCTGGACGGTGGACGGCGCCAGAGCGCGAGCGTGCCCGTCGAC  
CTGGTGCTCGGCAGCTTCCATGCGAAGTGGGAGGTGATGGAGCGCCTCGC  
GCAGGCGTACATCATCGGCACTCTCCGCATATGGAACGTCTTCTGCGCTG  
CTGGAGAGCGTCACACGATAGACGAGTTGCTCGTCAGGCTTCAAATCTCT  
GTCGTCTACAGGAAGGTCATCAAGCGATGGATGGAACACCTTGTCCGAT  
CGGCATCCTTGTAGGGGACGGAGAGCATTTTGTGAGCTCTCAGCCGCTGC  
CGGAGCCTGATTTGGCGGCGGTGCTCGAGGAGGCCGGGAGGGTGTTCGCC  
GACCTCCCAGTCCTATTTGAGTGGTGCAAGTTTGCCGGGGAACGGCTCGC  
GGACGTATTGACCGGTAAGACGCTCGCGCTCGAGATCCTCTTCCCTGGTG  
GCTCGTTCGATATGGCGGAGCGAATCTATCGAGATTCGCCCATCGCCCGT  
TACTCGAACGGCATCGTGCGCGGTGTCGTCGAGTCGGCGGCGCGGGTGGT  
AGCACCGTCGGGAATGTTTACGATCTTGGAGATCGGAGCAGGGACGGGCG  
CGACCACCGCCGCGCTCCTCCCGGTGTTGCTGCCTGACCGGACGGAGTAC  
CATTTACCGATGTTTCTCCGCTCTTCCTTGCTCGCGCGGAGCAAAGATT  
TCGAGATTATCCATTCCTGAAGTATGGCATTCTGGATGTCGACCAGGAGC  
CAGCTGGCCAGGGATACGCACATCAGAGGTTTGACGTCATCGTCGCGGCC  
AATGTCATCCATGCGACCCGCGATATAAGAGCCACGGCGAAGCGTCTCCT  
GTCGTTGCTCGCGCCCGGAGGCCTTCTGGTGCTGGTCGAGGGCACAGGGC  
ATCCGATCTGGTTCGATATCACCACGGGATTGATTGAGGGGTGGCAGAAG

5

10

15

20

25

30

35

40

45

50

55

TACGAAGATGATCTTCGTATCGACCATCCGCTCCTGCCTGCTCGGACCTG  
GTGTGACGTCCTGCGCCGGGTAGGCTTTGCGGACGCCGTGAGTCTGCCAG  
GCGACGGATCTCCGGCGGGGATCCTCGGACAGCACGTGATCCTCTCGCGC  
GCGCCGGGCATAGCAGGAGCCGCTTGTGACAGCTCCGGTGAGTCGGCGAC  
5 CGAATCGCCGGCCGCGCGTGCAGTACGGCAGGAATGGGCCGATGGCTCCG  
CTGACGTCGTCCATCGGATGGCGTTGGAGAGGATGTACTTCCACCGCCGG  
CCGGGCCGGCAGGTTTGGGTCCACGGTCGATTGCGTACCGGTGGAGGCGC  
GTTACGAAGGCGCTCGCTGGAGATCTGCTCCTGTTTGAAGACACCGGGC  
AGGTCTGTCAGAGGTTTACAGGGCTCCGCCTGCCGCAGCTCGAGGCTTCT  
10 GCTTTCGCGCCGCGGGACCCGCGGGAAGAGTGGTTGTACGCTTTGGAATG  
GCAGCGCAAAGACCCCTATACCAGAGGCTCCGGCAGCCGCGTCTTCTTCCT  
CCGCGGGGGCTTGGCTCGTGCTGATGGACCAGGGCGGGACAGGCGCTGCG  
CTCGTATCGCTGCTGGAAGGGCGAGGCGAGGCGTGCCTGCGCGTCATCGC  
GGGTACGGCATAACGCTGCCTCGCGCCGGGGCTGTATCAAGTCGATCCGG  
15 CGCAGCCAGATGGCTTTCATACCCTGCTCCGCGATGCATTCGGCGAGGAC  
CGGATTTGTGCGCGGCTAGTGATATGTGGAGCCTTGATGCGACGGCAGC  
AGGGGAGAGGGCGACAGCGGAGTCGCTTCAGGCCGATCAACTCCTGGGGA  
GCCTGAGCGCGCTTCTCTGGTGCAGGCGCTGGTGCGCCGGAGGTGGCGC  
AACATGCCGCGGCTTTGGCTCTTGACCCGCGCCGTGCATGCGGTGGGCGC  
20 GGAGGACGCAGCGGCCTCGGTGGCGCAGGCGCCGGTGTGGGGCTCGGT  
GGACGCTCGCGCTCGAGCATCCAGAGCTGCGGTGCACGCTCGTGGACGTG  
AACCCGGCGCCGTCTCCAGAGGACGCAGCCGCACTGGCGGTGGAGCTCGG  
GGCGAGCGACAGAGAGGACCAGGTGCGATTGCGCTCGGATGGCCGCTACG  
TGGCGCGCCTCGTGCGGAGCTCCTTTTCCGGCAAGCCTGCTACGGATTGC  
40 GGCATCCGGGCGGACGGCAGCTATGTGATCACCGATGGCATGGGGAGAGT  
GGGGCTCTCGGTGCGCGCAATGGATGGTGATGCAGGGGGCCCGCCATGTGG  
TGCTCGTGGATCGCGGCGGCGCTTCCGAGGCATCCCGGGATGCCCTCCGG  
TCCATGGCCGAGGCTGGCGCGGAGGTGCAGATCGTGGAGGCCGACGTGGC  
45 TCGGCGCGACGATGTCGCTCGGCTCCTCTCGAAGATCGAACCGTCGATGC  
CGCCGCTTCGGGGGATCGTGTACGTCGGACGGGACCTTCCAGGGCGACTCC  
30



5

10

15

20

25

30

35

40

45

50

55

TCGATGCTGGAGCTGGATGCCCGTCGCTTCAAGGAGTGGATGTATCCCAA  
GGTGCTCGGAGCGTGGAACCTGCACGCGCTGACCAGGGATAGATCGCTGG  
ACTTCTTCGTCTGTATTCTCGGGCACCTCGCTTCTGGGCTTGCCAGGA  
CAGGGGAGCCGCGCCGCCGGTGACGCCTTCTTGGACGCCATCGCGCATCA  
5 CCGGTGCAAGGTGGGCCTTACAGCGATGAGCATCAACTGGGGATTGCTCT  
CCGAAGCATCATCGCCGGCGACCCGAACGACGGCGGAGCACGGCTCGAA  
15 TACCGGGGGATGGAAGGCCTCACGCTGGAGCAGGGAGCGGCGGCGCTCGG  
GCGCTTGCTCGCACGACCCAGGGCGCAGGTAGGGGTGATGCGGCTGAATC  
TGCGCCAGTGGTTGGAGTTXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
20 10 XXXXTATGGTATAACTTATTGATTATAATACAGTATACAAAGGTCCCTTT  
TCAGSGACCCCTTTCGTATGTTGTAGCTGATTTTATTTTCTTCTTTCTTT  
TGGGTGCTTTTAATAGCCATATAATGATGTTTCATATAAAGTTAAAAGTTT  
CCCCATCTGTCTATCTATAGCGTCATGTTTTTCAGGATTTCTTAATTTCT  
25 GCAGCAGTTTAGATATATATTTTCGACGGTAACTTTATTATTAATGAATTA  
15 TGGGATTGATACATTTCCCTTTGTCCAATCTAAATCTTTTCCTGATTTCAA  
TACCTCACTCTCATCTCTATTAATTATAAAAACCGAGCTTTCCATACGGAC  
CTGTTAAATATGATTGTAATTGTCTATACTCAGATGGACCTAATTCTTCA  
30 AAATTTTAGCATCAAAAACAACTTGTCTTGTTTATAGTCCTCCAATAC  
TCGTTTCCAAAAATCAGATTTGCCACCATGGTGCCTATAATATCTCGTC  
20 TCTGAACTGCGTTACCATTTGGATGGGACTTGATGTCTGTTAGGTGGGAT  
35 GCAAATACTATTCTTAGTGCGTCTAAACACCATTGTTCAAATTCAGTGGC  
ACCTTCATTTCCCTATCGGTATCTGATCTAGATGAGTAGTTATTTGACCTA  
TTGTTTTATTTCTAATGGCTGAATTATCTGAAATAATATTAATATCATAT  
TCGTCATTTATCTCCTCTGCTTCCTCTGGAGCAAGTGCATTACGATTTAG  
40 25 ATTCAAACCAAGCCAGTAACATGGATGAATTAATAATTTCTCATTACTTT  
CAAAACCTTTATCTGGAGTTCGCCCCGTATGGCAAATGAATAGGATGAG  
GTGTTTTTATCACGAATACCAACAAATCCTACACTATACAACTTTGGAG  
45 AATCCACTTGCCTTTAACAACGTATTTTCAGACGTTATTTTAGGATCTC  
CATTTTCTTCGATTAATTTCGAAAGATGCTTCTATTTTTTTTAAGCACGTA  
30 TAAACTGTTAATTCAGGTTCAATGCTACGAAATGCACTAGTTATAACCTG

5

10

5

15

20

10

25

15

30

20

35

40

25

45

30

50

55

TATTGAAGGAAAGATCTTCTGATACTCTTCCAGAGATCTTCAAGTCTGG  
CCATGGAAATTGACTTGGCTGCATATTCTAGGTCAGTGTTTATGATAGTT  
TCTCTATTCTCTCTGAATGCGGAAAAAAGCTTCATTCAACAATGATAG  
TAAATCCCTGGGCCGGTAAAGGGTAAATTGCAAACATCGCTTAAACCAT  
TCCTCCCTTTAAGATCATCCGCTGTGCATCTATCCCAAACTCGTTGATCT  
TTCTCAATATCTAGCTTAAATGCTACTTTCATTCTTTTAGCTGACAGCAT  
TAGGAGTTGTGCCCAGTCCCAATGCAACCTTATGACTTGACCCTCTATAT  
TTCTCGAGTAATCAGGATCTTCCTTTGATAGCGACCTAAATATATTATCC  
CTTAAAAAATTATTGGACGAATGCATTTTGCTTTTGTATTAAATTCAAT  
AGATGCATATGCTAGACCTGCAATGATTCCAATTCCTATATTATCCGGTT  
CATACGCCTCATCTAGCTTATCCATTAATATGACAACTTTCCTGTCIGAG  
CGTTCAAGAAGTGATACTATATTATTTCTATTTCTGAGATATTCAAATT  
GAATTGAAGATCACCAATTGATTCTTCTGGGTTATTTTCATCTAAATACT  
CCTTTGCGACAAGCCTACACTTTCTTAAATGTACCTTGTGCAGAATTC  
CATTTTTTCAAATGTTCAATTCAACAATGTTTCTGATGATATTGAGATGA  
CAATTTGTAATGAGATGATATATATGATGCTATCTCCATTAGCATAGCGT  
ATCGCCATAATAGTCTTGTTGCTGCCCTTGCTAAATTAAATGATCCTGTA  
AATGGTTTCAACATTGATCTGAAACCAATAATTTGAGAATCGTCTGGTGA  
GAAACTCAGGATTAATATTTTTTTGTCTTTCTTCCAATGCTCATTTAGCT  
GAATAAATAAAGCACTTTTACCTGTCCCTCGTCTACCAACAACAATGGTC  
CTGTCTATCAGTTTCAATTAGAGTCCTAAAGTCAGCAGTTTCAATGAAAGC  
ATTACTCAACATCTTTTTATCATTTTCTGCTGTCTGATCACCACCGGAT  
TAGACTTTGAAGTAATATTCAATTCATATTCAACCTTTTATGTTAGTTG  
CTTTCATTTATTACTTTATATACTGTTGAACGAGCAATATTCATTGTTTT  
TGATATATGTGAGGCACCTAACCCTGTGCCACATATTTAATACTGCAT  
CTCTATCTATTTTTCTTTTTCTACCAAAAACAACCTCTTTTGCCA

SEQ ID No 82 (&gt;Contig57)

TCATCTATTGTATAGTTTGTATATTGATATGATATAATTATAACATATAA

ACAGTAACTTTCTCTACGTAGATCGAGGAGAAGACTCAATTTGTTGACA

5

10

15

20

25

30

35

40

45

50

55

TCXXXXXXXXXXXXXXXXXXXXXXXXXXXXXACTACTCGCATACCGTTG  
CGCAACAGCGGCGGAGGAGCAGGACGCATACGACATCACC GGCAATACG  
CTCAGCGTCGCCGACGGACGGTTGTCTTATACGCTAGGGCTGCAGGGACC  
CTGCCTGACCGTCGACACGGTCTGCTCGTCGCTCGTGGCCATCCACC  
5 TTGCCTGCCGACCGTGC GCGCTCGCGAGAGCGATCTCGCGCTGGCGGGA  
GGCGTCAACATGCTCCTTTCGTCCAAGACGATGATAATGCTGGGGCGCAT  
15 CCAGGCGCTGTCGCCCCGATGGCCACTGCCGGACATTGACGCCTCGGCCA  
ACGGGTTTCGTCCGTGGGGAGGGCTGCGGTATGGTTCGTGCTCAAACGGCTC  
TCCGACGCCCAGCGACACGGCGATCGGATCTGGGCTCTGATCCGGGGTTC  
20 GGCCATGAATCAGGATGGCCGGTCGACAGGGTTGATGGCACCCAATGTGC  
10 TCGCTCAGGAGGCGCTCTTGCGCGAGGCGCTGCAGAGCGCTCGCGTCGAC  
GCCGGGGCCATCGGTTATGTCGAGACCCACGGAACGGGGACCTCGCTCGG  
CGACCCGATCGAGGTCGAGGCGCTGCGTGCCGTGTTGGGGCCGGCGCGGG  
25 CCGATGGGAGCCGCTGCGTGCTGGGCGCAGTGAAGACAAACCTCGGCCAC  
15 CTGGAGGGCGCTGCAGGCGTGGCGGGTTTGATCAAGGCGGCGCTGGCTCT  
GCACCACGAACTGATCCCGCGAAACCTCCATTTCCACACGCTCAATCCGC  
30 GGATCCGGATCGAGGGGACCGCGCTCGCGCTGGCGACGGAGCCGGTGCCG  
TGGCCGCGGGCGGGCCGACCGCGCTTCGCGGGGTGAGCGCGTTCCGGCCT  
CAGCGGCACCAACGTCCATGTCGTGCTGGAGGAGGCGCCGGCCACGGTGC  
20 TCGCACCGGCGACGCCGGGGCGCTCAGCGGAGCTTTTGGTGCTGTGGCG  
35 AAGAGCGCCGCGCGCTGGACGCACAGGCGGCGCGGCTCTCAGCGCACAT  
CGCCGCGTACCCGAGCAGGGTCTCGGAGACGTGCGGTTTACGCTGGTAT  
CGACGCGTAGCCCGATGGAGCACCGGCTCGCGGTGGCGGCGACCTCGCGC  
40 GAGGCGCTGCGAAGCGCGCTGGAGGTTGCGGCGCAGGGGCAGACCCCGGC  
25 AGGCGCGGCGCGCGGCAGGGCCGCTTCCTCGCCCGGCAAGCTCGCCTTCC  
TGTTGCGCCGGGCAGGGCGCGCAGGTGCCGGGCATGGGCCGTGGGTGTGG  
GAGGCGTGGCCGGCGTTCCGCGAGACCTTCGACCGGTGCGTCACGCTCTT  
45 CGACCGGGAGCTCCATCAGCCGCTCTGCGAGGTGATGTGGGCCGAGCCGG  
GCAGCAGCAGGTGCTCGTTGCTGGACCAGACGGCGTTACCCAGCCGGCG  
30 CTCCTTGCGCTGGAGTACGCGCTGGCCGCGCTCTTCCGGTCGTGGGGCGT

5

10

15

20

25

30

35

40

45

50

55

GGAGCCGGAGCTCGTCGCTGGCCATAGCCTCGGCGAGCTGGTGGCCGCCT  
GCGTGGCGGGTGTGTTCTCCCTCGAGGACGCCGTGCGCTTGGTGGTCGCG  
CGCGGCCGGTTGATGCAGGCGCTGCCGGCCGGCGGCGGATGGTATCGAT  
CGCCGCGCCGGAGGCCGACGTGGCTGCCGCGGTGGCGCCGCACGCAGCGT  
5 TGGTGTGATCGCGGCAGTCAATGGCCCGAGCAGGTGGTGTGATCGCGGGC  
GCCGAGAAATTCGTGCAGCAGATC3CGGCGGCGTTTCGCGGCGCGGGGGGC  
15 GCGAACCAAACCGCTGCATGTCTCGCACGCGTTCCACTCGCCGCTCATGG  
ATCCGATGCTGGAGGCGTTCCGGCGGGTGAAGTGAAGTGAAGTGAAGTGAAGT  
CGGCCTTCGATCGCGCTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
20 TGAGGTGAGCGCGCCGGGTTACTG3GTGCGTCACGCGCGAGAGGCGGTGC  
10 GCTTCGCGGACGGAGTGAAGGCGCTGCACGCGGCGGCGGTGCGGGCCTCTTC  
GTCGAGGTGGGGCCGAAGCCGACGCTGCTCGGCCTTGTGCGGCCTGCCT  
GCCGGATGCCAGGCCGGTGTGCTCCAGCGTCGCGCGCCGGGCGTGACG  
25 AGGCTGCGAGCGCGCTAGAGGCGCTGGGTGGGTTCGCGGTGCTCGGTGGA  
15 TCGGTACCTGGTGGGTGTCTTCCCTTCGGGCGGACGGCGGGTACCGCT  
GCCAACCTATCCCTGGCAGCGCGAGCGTTACTGGATCGAAGCGCCGGTGC  
GACGACGCTGGACCGAAAGCGGCTGCCGTGGCTCGGCGAGCACCGGGCGC  
30 ATCGTGAGGCGGACGGCACCGGCCGTGCTCGGGCGGGGGGCCACCCCTT  
CTGGGTGAAGTCTTTTCCGTGTGACCCATGCCGTCTGCGCCTGTGGGA  
GACGACGCTGGACCGAAAGCGGCTGCCGTGGCTCGGCGAGCACCGGGCGC  
20 AGGGGGAGGTGCTGTTTCCCTGGCGCCGGGTACCTGGAGATGGCGCTGTCG  
35 TCGGGGGCCGAGATCTTGGGCGATGGACCGATCCAGGTACCGATGTGGT  
GCTCATCGAGACGCTGACCTTCGCGGGCGATACGGCGGTACCGGTCCAGG  
TGGTGACGACCGAGGAGCGACCGGGACGGCTGCGGTTCAGGTAGCGAGT  
40 CGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGCGT  
25 GCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTGGAACCTCGCCG  
CCCTGCGCGCCCGGCTTCATGCCGCGGTGCCCGCTGCGGCTATCTATGGT  
GCGCTCGCCGAGATGGGGCTTCAATACGGCCCGGCGTTGCGGGGGCTCGC  
45 CGAGCTGTGGCGGGGTGAGGGCGAGGCGCTGGGCAGGGTGAGACTGCCTG  
AGGCCGCCGGCTCCGCGACAGCCTACCAGCTGCATCCGGTGCTGCTGGAC  
30 GCGTGCGTCCAAATGATTGTTGGCGCGTTCCCGATCGCGATGAGGCGAC

5

10

15

20

25

30

35

40

45

50

55

GCCGTGGGCGCCGGTGGAGGTGGGCTCGGTGCGGCTGTTCCAGCGGTCTC  
CTGGGGAGCTATGGTGCCATGCGCGCGTCGTGAGCGATGGTCAACAGGCC  
TCCAGCCGGTGGAGCGCCGACTTTGAGTTGATGGACGGTACGGGCGCGGT  
GGTCGCCGAGATCTCCCGGCTGGTGGTGGAGCGGCTTGCGAGCGGTGTAC  
GCCGGCGCGACGCAGACGACTGGTTCCTGGAGCTGGATTGGGAGCCCCGCG  
GCGCTCGGTGGGCCCAAGATCACAGCCGGCCGGTGGCTGCTGCTCGGCGA  
GGGTGGTGGGCTCGGGCGCTCGTTGTGCTCGGCGCTGAAGGCCGCGGCC  
ATGTCGTCTGCCACGCCGCGGGGACGACACGAGCACTGCAGGAATGCGC  
GCGCTCCTGGCCAACGCGTTCGACGGCCAGGCCCCGACGGCCGTGGTGCA  
CCTCAGCAGCCTCGACGGGGGCGGCCAGCTCGGCCCGGGGCTCGGGGCGC  
AGGGCGCGCTCGACGCGCCCCGAGCCCAGATGTGATGCCGATGCCCTC  
GAATCGGCGCTGATGCGTGGTTCCGACAGCGTGCTCTCCCTGGTGCAAGC  
GCTGGTCCGCATGGACCTCCGAAACGCGCCGCGGCTGTGGCTCTTGACCC  
GCGGGGCTCAGGCGGCCGCGCCGCGCGATGTCTCCGTGGTGCAAGCGCCG  
CTGTTGGGGCTGGGCCGCACCATCGCCTTGGAGCACGCCGAGCTGCGCTG  
TATCAGCGTCGACCTCGATCCAGCCGAGCCTGAAGGGGAAGCCGATGCTT  
TGCTGGCCGAGCTACTTGACAGATGATGCCGAGGAGGAGGTGCGCTGCGC  
GGTGGCGACCGGCTCGTTGCGCGGCTCGTCCACCGGCTGCCCCACGCTCA  
GCGCCGGGAGAAGGTGAGCCCCGCGGTGACAGGCCGTTCCGGCTAGAGA  
TCGATGAACCCGCGCGCTGGACCAACTGGTGCTCCGAGCCACGGGGCGG  
CGCGCTCCTGGTCCGGGCGAGGTGAGATCTCCGTGCAAGCGGCGGGGCT  
CGACTCCATCGACATCCAGCTGGCGTTGGGCGTTGCTCCCAATGATCTGC  
CTGGAGAAGAAATCGAGCCGTTGGTGCTCGGAAGCGAGTGCGCCGGGCGC  
ATCGTCGCTGTGGGCGAGGGCGTGAACGGCCTTGTGGTGGGCCAGCCGGT  
GATCGCCCTTGCGGCGGGAGTATTTGCTACCCATGTCACCACGTGCGCCA  
CGCTGGTGTGCTCGGCCTCTGGGGCTCTCGGCGACCGAGGCGGCCGCG  
ATGCCCCCTCGCGTATTTGACGGCCTGGTACGCCCTCGACAAGGTCGCCCCA  
CCTGCAGGCGGGGAGCGGGTGCTGATCCATGCGGAGGCCGGTGGTGTGCG  
GTCTTTGCGCGGTGCGATGGGCGCAGCGCGTGGGCGCCGAGGTGTATGCG  
ACCGCCGACACGCCCGAGAACCGTGCCCTACCTGGAGTCGCTGGGCGTGCG

5

10

15

20

25

30

35

40

45

50

55

GTACGTGAGCGATTCCCGCTCGGGCCGGTTCGTCACAGACGTGCATGCAT  
GGACGGACGGCGAGGGTGTGGACGTCGTGCTCGACTCGCTTTCGGGCGAG  
CGCATCGACAAGAGCCTCATGGTCCTGCGCGCCTGTGGTCGCCTTGTGAA  
GCTGGGCAGGCGCGACGACTGCGCCGACACGCAGCCTGGGCTGCCGCCGC  
5 TCCTACGGAATTTTTCCTTCTCGCAGGTGGACTTGCGGGGAATGATGCTC  
GATCAACCGGCGAGGATCCGTGCGCTCCTCGACGAGCTGTTCCGGTTGGT  
CGCAGCCGGTGCCATCAGCCCACTGGGGTTCGGGGTTGCGCGTTGGCGGAT  
CCCTCACGCCACCGCCGGTCGAGACCTTCCCGATCTCTCGCGCAGCCGAG  
GCATTCCGGAGGATGGCGCAAGGACAGCATCTCGGGAAGCTCGTGCTCAC  
10 GCTGGACGACCCGGAGGTGCGGATCCGCGCTCCGGCCGAATCCAGCGTCG  
CCGTCCGCGCGGACGGCACCTACCTTGTGACCGGCGGTCTGGGTGGCCTC  
GGTCTGCGCGTGGCCGGATGGCTGGCCGAGCGGGGCGGGGGCAACTGGT  
GCTGGTGGGCCGCTCCGGTGGCGGAGCGCAGAGCAGCGAGCCGCCGTGG  
25 CGGCGCTGGAGGCCACGGCGCGCGCTCACGGTGGCGAAAGCGGACGTC  
GCCGATCGGTACAGATCGAGCGGGTCCTCCGCGAGGTTACCGCGTCGGG  
GATGCCGCTGCGGGGTGTCTGTCATGCGGCAGGTCTCGTGGATGACGGGC  
TGCTGATGCAGCAGACTCCGGCGCGGTTCCGCACGGTGATGGGACCTAAG  
GTCCAGGGGGCCTTGCACTTGACACCGCTGACACGCGAAGCGCCTCTTTC  
CTTCTTCGTGCTGTACGCTTCTGCAGCTGGGCTTTTCGGCTCGCCAGGCC  
20 AGGGCAACTATGCCGCAGCCAACGCGTTCCTCGACGCCCTTTCGCATCAC  
CGAAGGGCGCAGGGCCTGCCGGCGCTGAGCATCGACTGGGGCATGTTTAC  
GGAGGTGGGGATGGCCGTTGCGCAAGAAAACCGTGGCGCGCGGCAGATCT  
CTCGCGGGATGCGGGGCATCACCCCGATGAGGGTCTGTGAGCTCTGGCG  
CGCTTGCTCGAGGGTGATCGCGTGACAGCGGGGTGATACCGATCACTCC  
35 GCGGCAGTGGGTGGAGTTCTACCCGGCAACAGCGGCCTCACGGAGGTTGT  
CGCGGCTGGTGACCACGCAGCGCGCGGTGCTGATCGGACCGCCGGGGAT  
CGGGACCTGCTCGAACAGCTTGCGTCGGCTGAGCCGAGCGCGGGCGGG  
GCTGCTGCAGGACGTCGTGCGCGTGACAGGTCTCGCATGTGCTGCGTCTCC  
40 CTGAAGACAAGATCGAGGTGGATGCCCCGCTCTCGAGCATGGGCATGGAC  
CTGAAGACAAGATCGAGGTGGATGCCCCGCTCTCGAGCATGGGCATGGAC  
25 TCGCTGATGAGCCTGGAGCTGCGCAACCGCATCGAGGCTGCGCTGGGCGT

5

10

15

20

25

30

35

40

45

50

55

CGCCGCGCCTGCAGCCTTGGGGTGGACGTACCCAACGGTAGCAGCGATAA  
CGCGCTGGCTGCTCGACGACGCCCTCGTCGTCCGGCTTGGCGGCGGGTCCG  
GACACGGACGAATCGACGGCGAGCGCCGGTTCGTTCGTCCACGTCCTCCG  
CTTTCGTCTGTCTCAAGCCGCGGGCTCGTCTTCTGTCTTTCACGGTT  
5 CTGGCGGCTCGCCGAGGGCTTCCGTTCCTGGTCGGAGAAGTCTGAGTGG  
AGCGATCTGGAAATCGTGGCCATGTGGCAGATCGCAGCCTCGCCTCCGA  
GGACGCGCCTGGTAAGAAGTACGTCCAAGAGGCGGCCTCGCTGATTCAGC  
ACTATGCAGACGCACCGTTTGCCTTAGTAGGGTTCAGCCTGGGTGTCCGG  
TTCGTCTATGGGGACAGCCGTGGAGCTCGCCAGTCGTTCGGGCGCACCGGC  
10 TCCGCTGGCCGTCTTCACGTTGGGCGGCAGCTTGATCTCTTCTTCAGAGA  
TCACCCCGGAGATGGAGACCGATATAATAGCCAAGCTCTTCTTCCGAAAT  
GCCGCGGGTTTCGTGCGATCCACCCAACAAGTCCAGGCCGATGCTCGCGC  
AGACAAGGTCATCACAGACACCATGGTGGCTCCGGCCCCCGGGGACTCGA  
25 AGGAGCCGCCCCGTGAAGATCGCGGTCCCTATCGTCGCCATCGCCGGCTCG  
GACGATGTGATCGTGCCTCCGAGCGACGTTCAAGATCTACAATCTCGCAC  
CACGGAGCGCTTCTATATGCATCTCCTTCCCGGAGATCAGGAATTTCTCG  
TCGATCGAGGGCGCGAGATCATGCACATCGTCGACTCGCATCTCAATCCG  
CTGCTCGCCGCGAGGACGACGTCGTCAAGCCCCGCTTCGAGGCAAAATG  
ATGGCAGCCTCCCTCGGGCGCGGAGATGGTTGGGAGCAGCGTGGGCGCT  
20 GGCGGCCGCGCGCAGGCCGCGGAGGCGCATGAGCCTTCTTGGACGTTTGC  
AGTATAGGAGATTTTATGACACAGGAGCAAGCGAATCAGAGTGAGACGAA  
GCCTGCTTTTCGACTTCAAGCCGTTTCGCGCCTGGGTACGCGGAGGACCCGT  
TCCCCGCGATCGAGCGCCTGAGAGAGGCAACCCCCATCTTCTACTGGGAT  
GAAGGCCGCTCCTGGGTCTCTACCCGATACACGACGTGTGGCGGTGTT  
40 CCGCGACGAACGCTTCGCGGTCACTCGAGAAGAGTGGGAATCGAGCGCGG  
AGTACTCGTCGGCCATTCCCGAGCTCAGCGATATGAAGAAGTACGGATTG  
TTCGGGCTGCCGCCGAGGATCACGCTCGGGTCCGCAAGCTCGTCAACCC  
45 GTCGTTTACGTACGCGCCATCGACCTGCTGCGCGCCGAAATACAGCGCA  
CCGTGACACGCTGCTCGATGCTCGCTCCGGACAAGAGGAGTTCGACGTT  
30 GTGCGGGATTACGCGGAGGGAATCCCGATGCGCGCGATCAGCGCTCTGTT

5

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GAAGGTTCCGGCCGAGTGTGACGAGAAGTTCCGTCGCTTCGGCTCGGCGA  
CTGCGCGCGCGCTCGGCGTGGGTTTGGTGCCCCAGGTCGATGAGGAGACC  
AAGACCCTGGTCGCGTCCGTCACCGAGGGGCTCGCGCTGCTCCATGACGT  
CCTCGATGAGCGGCGCAGGAACCCGCTCGAAAATGACGTCTTGACGATGC  
TGCTTCAGGCCGAGGCCGACGGCAGCAGGCTGAGCACGAAGGAGCTGGTC  
GCGCTCGTGGGTGCGATTATCGCTGCTGGCACCGATAACCACGATCTACCT  
TATCGCGTTTCGCTGTGCTCAACCTGCTGCGGTCGCCCCAGGCGCTCGAGC  
TGGTGAAGGCCGAGCCCCGGGCTCATGAGGAACGCGCTCGATGAGGTGCTC  
CGCTTCGACAATATCCTCAGAATAGGAACTGTGCGTTTCGCCAGGCAGGA  
CCTGGAGTACTGCGGGGCATCGATCAAGAAAGGGGAGATGGTCTTTCTCC  
TGATCCCGAGCGCCCTGAGAGATGGGACTGTATTCTCCAGGCCAGACGTG  
TTTGATGTGCGACGGGACACGGGCGCGAGCCTCGCGTACGGTAGAGGCC  
CCATGTCTGCCCCGGGGTGTCCCTTGCTCGCCTCGAGGCGGAGATCGCCG  
TGGGCACCATCTTCCGTAGGTTCCCCGAGATGAAGCTGAAAGAAACTCCC  
GTGTTTGGATACCACCCCGCGTTCCGGAACATCGAATCACTCAACGTCAT  
CTTGAAGCCCTCCAAAGCTGGATAGCTCGCGGGGGTATCGCTTCCCGAAC  
CTCATTCCCTCATGATACAGCTCGCGCGCGGGTGCTGTCTGCCGCGGGTG  
CGATTTCGATCCAGCGGACAAGCCCATGTGTCAGCGCGCGAAGATCGAATCC  
ACGGCCCCGAGAAGAGCCCGTCCGGGTGACGTCCGAAGAAGTGCCGGGCG  
CCGCCCTGGGAGCGCAAAGCTCGCTCGTTCGCGCTCAGCACGCCGCTCGT  
CATGTCCGGCCCTGCACCCGCGCCGAGGAGCCGCCCGCCCTGATGCACGG  
CCTCACCGAGCGGCAGGTTCTGCTCTCGCTCGTCCGCTCGCGCTCGTCC  
TCCTGACCGCGCGCGCCTTCGGCGAGCTCGCGCGGCGGCTGCGCCAGCCC  
GAGGTGCTCGCGGAGCTCTTCGGCGGCGTGGTGCTGGGCCCCGTCCGTCTGT  
CGGCGCGCTCGCTCCTGGGTTCATCGAGTCCTCTTCCAGGATCCGGCGG  
TCGGGGTCGTGCTCTCCGGCATCTCCTGGATAGGCGCGCTCGTCTGTGCTG  
CTCATGGCGGGTATCGAGGTGATGTGAGCATCCTGCGCAAGGAGGCGCG  
CCCCGGGGCGCTCTCGGCGCTCGGCGCGATCGCGCCCCCGCTGCGCACGC  
CGGGGCCGCTGGTGCAGCGCATGCAGGGCGCGTTACGTGGGATCTCGAC  
GTCTCGCCGCGACGCTCTGCGCAAGCCTGAGCCTCGGCGCCTGCTCGTAC



5

10

15

20

25

30

35

40

45

50

55

ACCTCGCCGGTGCTCGCTCCGCCCGCGGACATCCGGCCGCCCGCCGCGGC  
CCAGCTCGAGCCGGACTCGCCGGATGACGAGGCCGACGAGGCCGACGAGG  
CGCTCCGCCCGTTCCGCGACGCGATCGCCGCGTACTCGGAGGCCGTTTCGG  
TGGGCGGAGGCCGCGCAGCGGCCGCGGCTGGAGAGCCTCGTGCGGCTCGC  
5 GATCGTGCGGCTGGGCAAGGCGCTCGACAAGGTCCCTTTCGCGCACACGA  
CGGCCGGCGTCTCCAGATCGCCGGCAGACTCCAGAACGATGCGGTCTGG  
15 TTCGATGTCGCCGCCCGGTACGCGAGCTTCCGCGCGGCGACGGAGCACGC  
GCTCCGCGACGCGGCGTCGGCCATGGAGGCGCTCGCGGCCGGCCCGTACC  
GCGGATCGAGCCGCGTGTCCGCTGCCGTAGGGGAGTTTCGGGGGGAGGCG  
20 GCGCGCCTTCACCCCGCGGACCGTGTACCCGCGTCCGACCAGCAGATCCT  
10 GACCGCGCTGCGCGCAGCCGAGCGGGCGCTCATCGCGCTCTACACTGCGT  
TCGCCCCGTGAGGAGTGAGCCTCTCTCGGGCGCAGCCGAGCGGCGGCGTGC  
CGGTGGTTCCCTCTTCGCAACCATGACCGGAGCCGCGCTCGGTCCGCGCA  
25 GCGGCTAGCGCGCGTCGCGGCAGAGATCGCTGGAGCGACAGGCGACGACC  
15 CGCCCGAGGGTGTCTGAACGGATTGCCGCGAGCCCTCATTGCGGATCCCCTC  
CAGACACTCGTTCAGCTGCTTGGCGTCGATGCCGCTGGGCACTCGCCGA  
30 AGGTCAGCTCGTCGCGCCACTCGGATCGGATCTTGTTTCGAGCACGCGTCC  
TTGCTCGAATACTCCCGGTCTTGTTCCGATGTTGTTGCACCGCGCCTCGCG  
GTCGCACCGCGCCGCCACGATGCTATCGACGGCGCTGCCGACTGGCACCG  
20 GCGCCTCGCCCTGCGCGCCACCCGGGGTTTGCGCCTCCCCGCCTGACCGC  
35 TTTTCGCCGCGCGCACGCCGCGAGCAGGCTCATTCCCGACACCGAGATCAG  
GCCCCAGACAGCTTCCCAGCAATCTTTTGATGGCTTCCCCTCCCTCAC  
GACACGTCACATCAGAGACTCTCCGCTCGGCTCGTCGGTTCGACAGCCGG  
40 CGACGGCCACGAGCAGAACCGTCCCCGACCAGAACAGCCGCATGCGGGTT  
25 TCTCGCAACATGCCCCGACATCCTTGCGACTAGCGTGCCCTCCGCTCGTGC  
CGAGATCGGCTGTCTGTGCGACGGCAATATCCTGCGATCGGCCGGGCAG  
GAGGTACCGACACGGGCGCCGGGCGGGAGGTGCCGCCACGGGCTCGAAAT  
45 GTGCTGCGGCAGGCGCCTCCATGCCCGCAGCCGGAACGCGGCGCCCGGC  
CAGCCTCGGGGTGACGCCGCAAACGGGAGATGCTCCCGGAGAGGCGCCGG  
30 GCACAGCCGAGCGCCGTCACCACCGTGCGCACTCGTGAGCTCCAGCTCCT

5

10

15

20

25

30

35

40

45

50

55

CGGCATAGAAGAGACCGTCACTCCCGGTCCGTGTAGGCCATCGTGCTGAT  
CAGCGCGTTCTCCGCCTGACGCGAGTCGAGCCGGGTATGCTGCACGACAA  
TGGGAACGTCCGATTTCGATCACGCTGGCATAGTCCGTATCGCGCGGGATC  
GGCTCGGGTTCGGTCAGATCGTTGAACCGGACGTGCCGGGTGCGCCTCGC  
TGGGACGGTCACCCGGTACGGCCCCGGCGGGTTCGCGGTGCTGAAGTAGA  
CGGTGATGGCGACCTGCGCGTCCCGGTCCGACGCATTCAACAGGCAGGCC  
GTCTCATGGCTCGTCATCTGCGGCTCGGGTCCGTTGCTCCGGCCTGGGAT  
GTAGCCCTCTGCGATTGCCAGCGCGTCCGCCCGATCGGCTTCTCCATAT  
GTCCTCCCTGCTGGCTCCTCTTTGGCTGCCTCCCTCTGCTGTCCAGGAGC  
GACGGCCTCTTCTCCCGACGCGCTCGGGGATCCATGGCTGAGGATCCTCG  
CCGAGCGCTCCTTGCCGACCGGCGCGCCGAGCGCCGACGGGCTTTGAAAG  
CACGCGACCGGACACGTGATGCCGGCGCGACGAGGCCGCCCGCGTCTGA  
TCCCGATCGTGACATCGCGACGTCCGCCGGCGCCTCTGCAGGCCGGCCTG  
AGCGTTGCGCGGTTCATGGTCGTCTCGCGTCACCGCCACCCGCCGATTCA  
CATCCCACCGCGGCACGACGCTTGCTCAAACCGCGGCGAGACGGCCGGGC  
GGCTGTGGTACCGGCCAGCCCGGACGCGAGGCCCGAGAGGGACAGTGGGT  
CCGCCGTGAAGCAGTGAGGCGATCGAGGTGGCAGATGAAACACGTTGACA  
CGGGCCGACGAGTCGGCCGCGGATAGGGCTCACGCTCGGTCTCCTCGCG  
AGCATGGCGCTCGCCGGCTGTGGCGGCCCGAGCGAGAAAATCGTGCAAGG  
CACGCGGCTCGCGCCCCGGCGCCGATGCGCACGTCCGCCCGACGTGACC  
CCGACGCCGCGACCACGCGGCTGGCGGTGGACGTGTTACCTCTCGCCG  
CCCGAGCGCATCGAGGCCGGCAGCGAGCGGTTTCGTCTGTCAGCGTCC  
GAGCTCCGAGTCCCCGTGGCAACGGGTTCGGAGTGCTCGACTACAACGCTG  
CCAGCCGAAGAGGCAAGCTGGCCGAGACGACCGTGCCGCATGCCAACTTC  
GAGCTGCTCATACCGTCGAGAAGCAGAGCAGCCCTCAGTCTCCATCTTC  
TGCCGCCGTTCATCGGGCCGACGTCCGTTCGGGTAACATCGCGCTATCAGCA  
GCGCTGAGCCCGCCAGCAGGCCCCAGAGCCCTGCCTCGATCGCCTTCTCC  
ATCATATCATCCCTGCGTACTCCTCCAGCGACGGCCGCGTCGAAGCAACC  
GCCGTGCCGGCGCGGCTCTACGTGCGCGACAGGAGAGCGTCCTGGCGCGG  
CCTGCGCATCGCTGGAAGGATCGGCGGAGCATGGAGAAAGAATCGAGGAT

5

10

15

20

25

30

35

40

45

50

55

CGCGATCTACGGCGCCATCGCAGCCAACGTGGCGATCGCGGCGGTCAAGT  
TCATCGCCGCCCGCGTGACCGGCAGCTCGGCGATGCTCTCCGAGGGCGTG  
CACTCCCTCGTCGATACTGCAGACGGGCTCCTCCTCCTGCTCGGCAAGCA  
CCGGAGCGCACGCCCCGCCGACGCCGAGCATCCGTTCCGGCCACGGCAAGG  
5 AGCTCTATTTCTGGACGCTGATCGTCGCCATCATGATCTTCGCCGCGGGC  
GGCGGCGTCTCGATCTACGAAGGGATCTTGACCTCTTGACCCGCGCCA  
15 GATCGAGGATCCGACGTGGAACACTACGTGCTCCTCGGCGCAGCGGCCGTCT  
TCGAGGGGACGTGCTCATCATCTCGATCCACGAGTTCAAGAAGAAGGAC  
GGACAGGGCTACCTCGCGGCGATGCGGTCCAGCAAGGACCCGACGACGTT  
20 CACGATCGTCCTGGAGGACTCCGCGGCGCTCGCCGGGCTCACCATCGCCT  
TCCTCGGCGTCTGGCTCGGGCACCCCTGGGAAACCCCTACCTCGACGGC  
GCGGCGTCGATCGGCATCGGCCTCGTGCTCGCCGCGGTGCGGGTCTTCCT  
25 CGCCAGCCAGAGCCGTGGGCTCCTCGTGGGGAGAGCGCGGACAGGGAGC  
TCCTCGCCGCGATCCGCGCGCTCGCCAGCGCAGATCCTGGCGTGTGCGCG  
15 GTGGGGCGGCCCCCTGACGATGCACTTCGGTCCGCACGAAGTCCTGGTCGT  
GCTGCGCATCGAGTTGACGCGCGGCTCACGGCGTCCGGGGTCCGCGGAGG  
CGAGGGAGCGCATCGAGACCCGGATACGGAGCGAGCGACCCGACGTGAAG  
30 CACATCTACGTGAGGCCAGGTGCTCCACCAGCGCGCGAGGGCGTGACG  
CGCCGTGGAGAGACCGCGCGCGGCTCCGCCATCCTCCGCGGCGCCCCGG  
20 CTCAGGTGGCCCTCGCAGCAGGGCGCGCCTGGCGGGCAAACCGTGCAGAC  
35 GTCGTCCTTCGACGCGAGGTACGCTGGTTGCAAGTCGTCACGCCGTATCG  
CGAGGTCCGGCAGCGCCGGAGCCCGGGCGGGCCGGGCGCACGAAGGCGCG  
GCGAGCGCAGGCTTCGAGGGGGGCGACGTCATGAGGAAGGCCAGGGCGCA  
40 TGGGGCGATGCTCGGCGGGCGAGATGACGGCTGGCGTCCGGCCTCCCCG  
25 GCGCCGGCGCGCTTCGCGCCGCGCTCCAGCGCGGTGCTCGCGCGATCTC  
GCCCCGGCGCCGGCTCATCGCCTCCGTGTCCCTCGCCGGCGGCGCCAGCAT  
GGCGGTGCTCTCGCTGTTCCAGCTCGGGATCATCGAGCGCCTGCCCGATC  
45 CTCCGCTTCCAGGGTTCGATTGCGCCAAGGTGACGAGCTCCGATATCGCG  
TTCGGGCTCACGATGCCGGACGCGCCGCTCGCGCTCACCAGCTTCGCGTC  
30 CAACCTCGCGCTGGCTGGCTGGGGAGGCGCCGAGCGCGCCAGGAACACCC

5

10

15

20

25

30

35

40

45

50

55

CCTGGATCCCCGTGCGCGTGGCGGCCAAGGCGGCCGTCGAGGCGGCCGTG  
TCCGGATGGCTCCTCGTCCAGATGCGACGGCGGGAGAGGGCCTGGTGCGC  
GTACTGCCTGGTCGCCATGGCGGCCAACATGGCCGTGTTTCGCGCTCTCGC  
TCCCGGAAGGGTGGGCGGCGCTGGGGAAGGCGCGAGCGCGCTCGTGACAG  
5 GACGGGCGCGGGCAGCCCCGGCCATCGGAGGCCGGCGTGACCCGCTCCG  
TCACGCCCCAGCCCGCGCCGCGTGATCTCCCGCGGACAGGGCGCGTACCG  
TGGACCCCGCACGCGCCGCGTCGACGGACATCCCCGGCGACCCGCGCGGC  
GCGACCCGCGCAACTCCGGCCCCGCCGCGGGCATCGACATCTCCCGTGAG  
CAAGGGCACTCCGCTCCTGCCCCGCGTCCGCGAACGATGGCTGCGCTGTTT  
10 CCACCCTGGAGCAACTCCGTTTACCGCGTGGCGCTCGTCGGGCTCGTCGC  
CTCGGCGGGCGGCGCCATCCTCGCGCTCATGATCTACGTCCGCACGCCGT  
GGAAGCGATACCAGTTTCGAGCCCGTCGATCAGCCGGTGCAGTTTCGATCAC  
CGCCATCACGTGCAGGACGACGCCATCGATTGCGTCTACTGCCACACCAC  
25 GGTGACCCGCTCGCCACGGCGGGGATGCCGCCGACGGCCACGTGCATGG  
GGTGCCACAGCCAGATCTGGAATCAGAGCGTCATGCTCGAGCCCGTGCGG  
15 CGGAGCTGGTTCTCCGGCCACGCCGATCCCGTGGAACCGGGTGAAACTCC  
GTGCCCCACTTCGTCTATTTCAACCACGCGATCCACGTGAACAAGGGCGT  
GGGCTGGCGTGAAGCTGCCACGGGCGCGTGGACGAGATGGCGGCCGTCTA  
CAAGGTGGCGCCGATGACGATGGGCTGGTGCTGGAGTGCCATCGCCTGC  
20 CGGAGCCGCACCTCCGCCCCGCTCTCCGCGATCACCGACATGCGCTGGGAC  
35 CCGGGGGAGCGGAGGGATGAGCTCGGGGCGCAGCTCGCGAAGGAATACGG  
GGTCCGGCGGCTCACGCACTGCACAGCGTGCCATCGATGAACGATGAACA  
GGGGATCTCCTTGAAAGACGCAGATGAGATGAAGGAATGGTGGCTAGAAG  
CGCTCGGGCCGGCGGGAGAGCGCGCTCCTACAGGCTGCTGGCGCCGCTC  
40 ATCGAGAGCCCGGAGCTCCGCGCGCTCGCCGCGGGCGAACC GCCCCGGGG  
25 CGTGGACGAGCCGGCGGGCGTCAGCCGCCGCGCGCTGCTCAAGCTGCTCG  
GCGCGAGCATGGCGCTCGCCGGCGTCGCGGGCTGCACCCCGCATGAGCCC  
45 GAGAAGATCCTGCCGTACAACGAGACCCCGCCGCGTCGTGCCGGGTCT  
CTCCCAATCCTACGCGACGAGCATGGTGCTCGACGGGTATGCCATGGGCC  
30 TCCTCGCCAAGAGCTACGCGGGGCGGCCCATCAAGATCGAGGGCAACCCC

180

5

10

15

20

25

30

35

40

45

50

55

GCGCACCCGGCGAGCCTCGGCGCGACCGGCGTCCACGAGCAGGCCTCGAT  
CCTCTCGCTGTACGACCCGTACCGCGCGCGCGCCGACGCGCGGCGGCC  
AGGTCGCGTCGTGGGAGGCGCTCTCCGCGCGCTTCGGCGGCGACCGCGAG  
GACGGCGGCGCTGGCCTCCGCTTCGTCTCCAGCCCACGAGCTCGCCCCCT  
5 CATCGCCGCGCTGATCGAGCGCGTCCGGCGCAGGTTCCCCGGCGCGCGGT  
TCACCTTCTGGTCGCCGGTCCACGCCGAGCACGCGCTCGAAGGCGCGCGG  
GCGGCGCTCGGCCTCAGGCTCTTGCCCTCAGCTCGACGTCGACCAGGCCGA  
GGTGATCCTCTCGCTGGACGCGGACTTCCTCGCGGACATGCCGTTACGCG  
TGCGCTATGCGCGCGACTTCGCCGCGCGCCGCCGCCCGCGAGCCCGGCG  
10 GCGGCCATGAGCCGCCTCTACGTCGCGGAGGCGATGTTACGCCCCACGGG  
GACGCTCGCCGACCACCGGCTCCGCGTGCGGCCCCGCCGAGGTCGCGCGCG  
TCGCGGCCGGCGTCGCGGTGGGAACCTCGTGACGAGTCTTTGTCTTGCGC  
CCTGTCCGGGAATAACGGACACCTTATCGCGGGTCGCTCTTTGTGCGCGG  
25 CTTCTGTACCTCTCAGGACAGGTAGAAGAGGGACTCAGGGGCCCTTATGT  
15 TAACTGGGGATGCCTTCGGGACGGCCGCAAATATATCCTATCACCTCACT  
GGGTGTGGGGGAGCACCGCGAGGATGTACAACCTCTGTAACCTCTATGTGA  
GATAATGTGTGCAGTGATCTGAGACTTATTTGTGTGACCGAGACGTCTCT  
CTTATTGGTACGCATAGTATAATATAACACGTCTCATACTACTCCCGAC  
ATATCCGCGGTATGCGCGCACATAGAATAGGTGATGATAAATCCCTAGTG  
20 TGTGGAAGTAGAAGATGCGGGAGTTACCTGATATTTACGGAAAAAGTATT  
35 ATCTCAACTACCTCTCTGTTGAGACTATCACTTCGGTGTCGTTGTGCTGC  
TGGT,

or its complementary strand,

25

(b) DNA-sequences which hybridise under stringent conditions  
to regions of DNA-sequences according to (a) encoding proteins  
or to fragments of said DNA-sequences,

5

10

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

15

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

20

10. 11. Peptide encoded by a DNA sequence according to claim 10 selected from the group consisting of

25

Seq ID No 83

>Contig56\_003 2890 amino acids MW=307428 D pI=5.76 numambig=13

15 IRPRAAAVPMRSTVTGGVIAGPELGASYWADNLRQPVRFAAAAQALLEGGPALFIEMSPH  
PILVPPLDEIQTAAEQGGAAVGSLRRQDERATLLEALGTLWASGYPSWARLFPAGGRR  
30 VPLPTYFWQHERCWIIEVEPDARRLAAADPTKDWFYRTDWPEVPRAAPKSETAHGSWLLLA  
DRGGVGEAVAAALSTRGLSCTVLHASADASTVAEQVSEAASRRNDWQGVLYLWGLDAVVD  
AGASADEVSEATTRATAPVLGLVRFSLAAPHPPRFWVVTGACTVGGEPEASLCQAALWG  
20 LARVAALEHPAAWGGLVDLDPQKSPTIEPLVAELLSFDAEDQLAFRSGRRHAARLVAAP  
35 PEGDVAPISLSAEGSYLVTGGLGGLGLLVARWLVERGARHLVLTSRHGLPERQASGGEQP  
PEARARIAAVEGLEAQGARVTVAADVVAEADPMTALLAAIEPPLRGVVHAAGVFPVRHLA  
ETDEALLESVLRPKVAGSWLLHRLLRDRPLDLFVLFSSGAAVWGGKGQAYAAANAFLDG  
40 LAHHRRRAHSLPALSLAWGLWAEGMVDAKAHARLSDIGVLPMTGPALSALERLVNTSAV  
25 QRSVTRMDWARFAPVYAARGRRNLLSALVAEDERAASPPVPTANRIWRGLSVAESRSALY  
ELVRGIVARVLGFSDFGALDVGRGFAEQGLDSLMALEIRNRLQRELGERLSATLAFDHPT  
VERLVHLLTDVLKLEDSDTRHIRSVAADDDIAIVGAACRFPGGDEGLETYWRHLAEGM  
45 VVSTEVPADRWRAADWYDPDEVFGRTYVAKGAFLRDVRSLDAAFFAISPREMSLDPQQ  
RLLLEVSWEAIERAGQDPMALRESATGVFVGMIGSEHAERVQGLDDDAALLYGTTGNLLS  
30 VAAGRLSFFLGLHGPTMTVDTACSSSLVALHLACQSLRLGECDAQALAGGSSVLLSPRSFV

50

55

5

10

15

20

25

15

30

35

40

45

30

50

55

AASRMRLSPDGRCKTFSAAADGFARAEGCAVVVLKRLRDAQDRDPILAVVRSTAINHD  
GPSSGLTVPSGPAQQALLRQALAQAGVAPAEVDFVECHGTGTALGDPFIEVQALGAVYGRG  
RPAERPLWLGAVKANLGHLEAAAGLAGVLKVLLEHEQIPAQPELDELNPHIPWAEPLV  
AVVRRAVPWPRGARPRRAGVSAFGLSGTNAHVLEEAPAVEPVAAAPERAAELFVLSAKS  
5 AAALDAQAARLRDHLEKHVELGLGDVAFSLATTSAMEHRLAVAASSREALRGALSAAAO  
GHTPPGAVRGRASGGSAPKVVFVFPQGSQWVGMGRKLMEEPVFRAALEGCDRAIEAEA  
15 GWSLLGELSADEAASQLGRIDVVQPVLFAMEVALSALWRSWGVEPEAVVGHSMEVAAA  
VAGALSLEDAVAIIICRRSRLRLRISGQGEMALVELSLEEAEALRGHEGRLSVAVSNSPR  
STVLAGEPAALSEVLAALTAKGVFWRQVKVDVASHSPQVDPLREELIAALGAIRPRAAV  
20 PMRSTVTGGVIAGPELGASYWADNLRQPVRFAAAAQALLEGGPALFIEMSPHPILVPPLD  
EIQTAEEQGGAAVGSLLRGQDERATLLEALGTLWASGYPVSWARLFPAGGRRVPLPTYPW  
QHRYWIEDSVHGSKPSLRLRQLRNGATDHPLLGAPLLVSARPGAHLWEQALSDERLSYL  
SEHRVHGEAVLPSAAYVEMALAAGVDLYGTATLVLEQLALERALAVPSEGGRIVQValse  
25 EGPGRASFQVSSREEAGRSWVRHATGHVCSGQSSAVGALKEAPWEIQRRCPSVLSSEALY  
PLLNEHALDYGPCFQGVQVWLGTGEVLGRVRLPGDMASSSGAYRIHPALLDACFQVLTA  
LLTTPESIEIRRLTDLHEPDLPRSRAPVNQAVSDTWLWDAALDGGRRQSASVPVDLVLG  
SFHAKWEVMERLAQAYIIGTLRIWNVFCAGGERHTIDELLVRLQISVVYRKVIKRWMEHL  
30 VAIGILVGDGEHFVSSQPLPEPDLAAVLEEAGRVFADLPVLFEWCKFAGERLADVLTGKT  
LALAILFPGGSFDMAERIYRDSPIARYSNGIVRGVVESAARVVAPSGMFSILEIGAGTGA  
TTAAVLPVLLPDRTEYHFTDVSPFLARAEQRFRDYPFLKYGILDVDQEPAGQGYAHQRF  
20 DVIVAANVIHATRDIRATAKRLLSLLAPGGLLVLEGTGHPIWFDITTGLIEGWQKYEDD  
LRIDHPLLPARTWCDVLRVGFADAVSLPGDGSPAGILGQHVLRSRAPGIAGAACDSSGE  
SATESPAARAVRQEWADGSADVHRMALERMYFHRRPGRQVWVHGRLRTGGGAFTKALAG  
35 DLLLFEDTGQVVAEVQGLRLPQLEASAFAPRDPREEWLYALEWQRKDPIPEAPAAASSSS  
AGAWLVLMDOGGTGAALVSLLEGRGEACVRVIAGTAYACLAPGLYQVDPAQPDGFHTLLR  
DAFGEDRICRAVVHMWSLDATEAGERATAESLQADQLLGSLSALSIVQALVRRRWNRNMPR  
LWLLTRAVHAVGAEDAAASVAQAPVWGLGRTLALHPELRCTLVDVNPAPSPEDAAALAV  
45 ELGASDREDQVALRSDGRYVARLVRSSFSGKPATDCGIRADGSYVITDGMGRVGLSVAQW  
MVMQGARHVVLVDRGGASEASRDALRSMAEAGAEVQIVEADVARRDDVARLLSKIEPSMP  
30 PLRGIVYVDGTFQGDSSMLELDARRFKEWMPKVLGAWNHALTRDRSLDFFVLYSSGTS

5

183

10

LLGLPGQGSRAAGDAFLDAIAHHRCKVGLTAMSINWGLLSEASSPATPNDGGARLEYRGM  
EGLTLEQGAAALGRLLARPRAQVGVMRLNLRQWLEXXXXXXXXXXXXXXWYNLLIIIIQYTK  
VPFQGPFRML\*

5 Seq ID No 84

15

>Contig56\_027 700 amino acids MW=80569 D pI=7.02 numambig=0  
MNMELNITSKSNPFGDTTAENDKKMLSNAFIETADFRTLIETDDRTIVVGRRGTKGSALF  
IQLNEHWKKDKKILILSFSPDDSQIIGFRSMLKPFTGSFNLARAATRLLWRYAMLMEIAS  
YISSHYKLSSQISSETLLNEHLKKWNSAQGDILRKCRVLAKEYLDENNPEESIGDLQFNL  
10 NISEIENNIVSLLERSDRKVVILMDKLDEAYEPDNIGIGIIAGLAYASIELNQKAKCIRP  
IIFLRDNIFRSLSKEDPDYSRNEGQVIRLHWDWAQLLMLSAKRMKVAFKLDIEKDQRVW  
DRCTADDLKGRNGFKRCLQFTLYRPRDLLSLLNEAFFSAFRENRETIINTDLEYAAKSIS  
25 MARLEDLWKEYQKIFPSIQVITSAFRSIEPELTVYTCLKKIEASFELIEENGDPKITSEI  
QLLKASGILQSLYSVGFVGIRDKNTSSYSFCHDGRTPDKGFESNEKLLIHPCYWLGLNLN  
15 RNALAPEEAEEINDEYDINIISDNSAIRNKTIGQITTHLDQIPIGNEGATEFEQWCLDAL  
RIVFASHLTDIKSHPNGNAVQRRDIIGTNGGKSDFWKRVLEDYKTRQVVFDAKNFEELGP  
30 SEYRQLQSYLTGPYKGLGFIINRDESEVLKSGKDLDTKEMYQSHNSLI IKLPKYISKL  
LQKLRNPEKHDAIDRQMGKLLTLYETSYMAIKSTQKKRRK\*

25

30

20 Seq ID No 85

35

>Contig57\_001 372 amino acids MW=38411 D pI=12.39 numambig=10  
MLTSXXXXXXXXXXLLAYRCATAARGAGRIRHHRQYAQRRRRTVVLYARAAGTLPDRRHG  
LLVVARGHPPCLPQPARSRERSRAGGRROHAPFVQDDDNAGAHPGA VARWPLPDIRRLGQ  
40 RVRPWGGLRYGRAQTALRRPATRRSDLGSDPGFGHESGWPVDRVDGTQCARS GGALARGA  
25 AERSRRRRGHRLCRDPRNGDLARRPDRGRGAACRVGAGAGRWEPLRAGRSEDKPRPPGGR  
CRRGGFDQGGAGSAPRTDPAKPPFFPHAQSADPDRGDRARAGDGAGAVAAGGPTALRGGER  
VRFQRHQRPCRAGGGAGHGARTGDAGALSGAFGAVGEERRRAGRTGGAALSAHRRVPGAG  
45 SRRRRVQPGIDA\*

50

55



5

Seq ID No 86

10

&gt;Contig57\_002 2259 amino acids MW=238258 D pI=5.92 numambig=0

MSYTLGLQGPCLTVDTVCSSSLVAIHLACRSLRARESDLALAGGVNMLLSSKTMIMLGRI

QALSPDGHCRTFDASANGFVRGEGCGMVVLKRLSDAQRHGDRIWALIRGSAMNQDGRSTG

5 LMAPNVLAQEALLREALQSARVDAGAIGYVETHGTGTSLGDPIEVEALRAVLGPARADGS

15

RCVLGAVKTNLGHLEGAAGVAGLIKAALALHHELI PRNLHFHTLNPRIRIEGTALALATE

PVPWPRAGRPRFAGVSAFGLSGTNVHVLEEAPATVLAPATPGRSAELLVLSAKSAAALD

AQAARLSAHIAAYPEQGLGDVAFSLVSTRSPMEHRLAVAATSREALRSALEVAAQGQTPA

GAARGRAASSPGKLAF LFAGQGAQVPGMGRGLWEAWPAFRETDFDRCVT LFDRELHQPLCE

20

10 VMWAE PGSSRSSLLDQTAFTQPALFALEYALAAALFRSWGVEPELVAGHSLGELVAACVAG

VFSLEDAVRLVVARGRLMQALPAGGAMVSI AAPEADVAAAVAPHAALVSIAAVNGPEQVV

IAGA EK FVQQIAAFAARGARTKPLHVSHAFHSPLMDPML EAFRRVTESVTYRRPSIALV

25

SNLSGKPCTDEVSAPGYWVRHAREAVRFADGVKALHAAGAGLFVEVGPKPTLLGLVPACL

PDARPVLLPASRAGRDEAASALEALGGFWVVGGSVTWSGVFPSSGRRVPLPTYPWQRERY

15 WIEAPVDREADGTGRARAGGHPLLGEVFSVSTHAGLRLWETTLD RKRLPWLGEHRAQGEV

VFPGAGYLEMALSSGAEILGDGP IQVTDVVL IETLTFAGDTAVPVQVVTTEERPGR LR FQ

30

VASREPGERRAPFRIHARGVLRRI GRVETPARSNLAALRARLHA AVPA AAIYGALAEMGL

QYGPALRGLAELWRGEGEALGRVRLPEAAGSATAYQLHPVLLDACVQMI VGAFADRDEAT

PWAPVEVGSVRLFQRSPGELWCHARVVS D GQQASSRWSADFELMDGTGAVVAEISRLVVE

20 RLASGVRRRDADDWFLELDWEPAALGGPKITAGRWLLGEGGGLGRSLCSALKAAGHV VV

35

HAAGDDTSTAGMRALLANAFD GQA PTAVVHLSSLDGGGQLGPGLGAQGALDAPRSPDVDA

DALESALMRGCD SVLSLVQALVGM DL RNAPRLWLLTRGAQAAAAGDVSVVQAPLLGLGRT

IALEHAELRCISVDLDPAEPEGEADALLAELLADDAEEVALRGGDRLVARLVHRLPDAQ

40

RREKVEPAGDRPFRLEIDEPGALDQLVLRATGRRAPGPGEVEISVEAAGLDSIDIQLALG

25 VAPNDLPGEEIEPLVLGSECAGRI VAVGEGVNLVVGQPVIALAAGVFATHVTTSATLVL

PRPLGLSATEAAAMPLAYLTAWYALDKVAHLQAGERVLIHAEAGGVGLCAVRWAQRVGAE

VYATADTPENRAYLES LGVRYVSDSRSGRFVTDVHAWTDGEGVDVVLDSLSGERIDKSLM

45

VLRACGRLVKLGRRDDCADTQPGLPPLLRNFSFSQVDLRGMMLDQPARIRALLDELFLV

AAGAISPLGSGLRVGGSLTPPPVETFPISRAAEAFRRMAQQQHLGKLVLTDDPEVRIRA

30 PAESSVAVRADGTYLVTGGLGGLGLRVAGWLAERGAGQLVLVGRSGAASAEQRAAVAAL

50

55

5

10

15

20

AHGARVTVAKADVADRSQIERVLREVTASGMPLRGVVHAAGLVDDGLLMQQTPARFRTVM  
GPKVQGALHLHTLTREAPLSFFVLYASAAGLFGSPGQGNIAAANAFLDALSHHRAQGLP  
ALSIDWGMFTEVGMAVAQENRGARQISRGMRGITPDEGLSALARLLEGDRVQTVIPITP  
RQWVEFYPTATAASRRLSRLVTTQRAVADRTAGDRDLLEQLASAEPSARAGLLQDVVRVQV  
5 SHVLRRLPEDKIEVDAPLSSMGMDLSLELRNRIEALGVAAPAALGWYPTVAAITRWL  
LDDALVVRLLGGGSDTDESTASAGSFVHVRFRPVVKPRARLFCFHGSGGSPGFRSWSEK  
SEWSDLEIVAMWHDRSLASEDAPGKKYVQEAASLIQHYADAPFALVGFSLGVRFVMGTAV  
ELASRSGAPAPLAVFTLGGSLISSSEITPEMETDIIAKLFFRNAAGFVRSTQQVQADARA  
DKVITDTMVAPAPGDSKEPPVKIAVPPIVAIAGSDDVIVPPSDVQDLQSRRTTERFYMHLLP  
10 GDHEFLVDRGREIMHIVDSHLNPLLAARTTSSGPAFEAK\*

Seq ID No 87

25

>Contig57\_027 419 amino acids MW=46737 D pI=5.09 numambig=0  
MTQEQANQSETKPAFDKPFAPGYAEDPFPPIERLREATPIFYWDEGRSWVLTRYHDVSA  
15 VFRDERFAVSREEWESSAEYSSAIPELSDMKKYGLFGLPPEDHARVRKLVNPSFTSRAID  
LLRAEIQRTVDQLLDARSGQEEFDVVRDYAEGIPMRAISALLKVPAECDEKFRFRFGSATA  
30 RALGVGLVPQVDEETKTLVASVTEGLALLHDVLDERRRNPLENDVLTMLLQAEADGSRLS  
TKELVALVGAIIAAGTDTTIYLIFAVLNLLRSPEALELVKAEPGLMRNALDEVLRFDNI  
LRIGTVRFARQDLEYCGASIKKGEMVFLLIIPALRDGTVFSRPDVFDVRRDTGASLAYGR  
20 GPHVCPGVSLARLEAEIAGVTIFRRFPKMLKETPVFGYHFAFRNIESLNVILKPSKAG\*

35

Seq ID No 88

40

>Contig57\_043 492 amino acids MW=52617 D pI=11.54 numambig=0  
MAARARKSCRARGSRPAPMRTSPPTSTPTPRPRGWRWTSFTSRRPSASRPAASGSSSGSV  
25 RAPSPRGNGSECSTTTLPAAEASWPRRPCRMPSTSSCSSPSRSRAALSLHLLPPSSGRRPS  
GNIALSAALSPPAGPRALPRSPSPSYHPCVLLQRRPRRSNRRAGAALRARQESVLARPAH  
RWKDRRSMEKESRIAIYGAIAANVAIAAVKFIAAAVTGSSAMLSEGVHSLVDADGLLLLL  
45 LGKHRSARPPDAEHPFGHGKELYFWTLIVAIMIFAAGGGVSIYEGILHLLHPRQIEDPTW  
NYVVLGAAAVFEGTSLIISIHEFKKKDGQGYLAAMRSSKDPTTFTIVLEDSAAAGLTIA  
30 FLGVWLGHRLGNPYLDGAASIGIGLVLA AVAVFLASQSRGLLVGESADRELLAAIRALAS

50

55

ADPGVSAVGRPLTMHFGPHEVLVVLRIEFDAALTASGVAEARERIETRIRSERPDVKHIY  
VEARSLHQRARA\*

12. DNA sequence according to any of claims 1 to 5 wherein the  
DNA is selected from the group consisting of

(a) the following DNA sequences:

Seq ID No 89 (>Contig10)

GGTAGTGAAATATGCTGTATTCAACAGAAAGCTTGATGAATTGATCTAGA  
AAGTAGAGCGAGAGAATCAAGTAAGATAGTAGGATGCATTATAAATATAG  
AATATATACTGCATACGATGACAGCATGCGCACGAATAGAATGCATAAGA  
GGCAAGCCAATAACCAAAAGTGGAGCCAGAGGAGATAGTCTCGCCAGTAG  
AAATAATGCTCAGCCAAGCGAGGTTGGACATATCAGTTCAGAGTAGGTC  
TCAACCCCGTATATGAGTCCAATGAAGCCTGTCTCATCCAGTTAACGGCC  
TTTTGAGCAGAGAATCCTCCCTATTTTCGGAGAGGACGCGTCGAATATAA  
AGCAGGTCCAAAGAAGCAAGCAATAGCCAAAAGTTTGAAAGGTTAGTACG  
AGCAGCGGCTGGAGGACACTATGGTCGTGCAACGGGGGTAAAGGGTTTCA  
CGTATTGTAGCAGAGCACGTCAGAGGGTTATTCGTGACATTGAGGCCAA  
CGAGGCGGTAGGACTTCGTAAGCGCATGACCATCCCGGTCACAAACGTAG  
TGCGGAGCGCCTCGTCACGCTCAACAAGGCCCTAGAACGCGCGGCGCAGA  
TCGACCCTTTTAAACGCCGGCACCAGCCGGACCGTCCTGCCAGGTTGT  
AAAGCGCTCCATCGGCCGACTTATGGCACTCGAGCCAAATCGCCCGGTTTC  
CCCATCGGTCAGCGCAAACGGCCCCCCCCGGGCGTCGCCACCCGCGGCGAC  
GAGGGGCGGTCCAGACGGGTGATCTCTCTCGTGAGCTCGCGGAGAGAGCC  
TCCTCGCAAGATCGATGTCAGCGGGATCGCGCGCCCCGTCCGCACCTGAA  
ACGCGTGCTGGAGCTCGACGGCAGCGAGGGAGTCGAGGCCGAACCGCGAT  
ATCGGCAGCGCGTCGTGATCTGCCCGGCGTCCAGACGAAGCGCGCGGGC  
GAGGGTCGAGCGCAGCGCGTCCAGCAGGCTCCGGCCGGAGGGCTCCTCGG  
TCTCCGGGGGCGCGTCGTCCGGGGGCGAGGCGTCGTGAGGAGCTCCGGC

5

10

15

20

25

30

35

40

45

50

55

GCGAACGCGACGTGGCGCTCGCCGAGCGCGTCCTCGAGAAAGGCGCGCCG  
GCACTCCCTCCGGCGGACCTTCCCGCTCGACGTCTTCGGCAGCGCGCCCG  
GCGCGATCAGCGCGACGGCGTGCGCGACGAGCTGGTGCTCGGCGGTCAAC  
GCCTCGCGCACGGCCGCCACGATCTCGCGCGGATCCGCGGCCACGCGCGG  
5 GTCGACCTCGCACACCACGGCGAGGCGCTCCTCGCCCTCGTGCTCCACGG  
AGAACGCGGCGCTGCAGCCCGGCCGACGGCGCGATGGCTGCTCTCGACG  
GTCTTCTCGATGTCCTGCGGGAAGTGGTTGCGGCCTCGAAGGATGATGAG  
GTCCTTCGACCTCCCCACCACGAACAGCTCGCCGCCCCGGAGGAAGCCGA  
GATCTCCCGTGCGCAGGTAGCGCGGCGCCGCGCTGCCAGCGAGCGTGCGC  
10 CCGAACGTGGCCTCCGTCTCCTCCGGGCGCCCCCAGTAGCCGACGGCTAC  
GCTGGGCCCCGGACACCCAGATCTCCCGATCTCCCCGGCCCCGAGCTCGT  
TCCCCGCGGGATCGACGATCGCGACCGCCCGCGGATCGAGCGCCCCGACCG  
CTGCCGACGAACACGCGCGCGCCCTCCGCGCCGACGCGACGGCGCGCCC  
25 GAGCTCCACCTCCTCGGGGGCGAGGCGCGCCAGCACCGGCGCCTCGGCCC  
15 GCGCTCCGCGCTCACGATGAGCGTGGCCTCGGCGAGCCCGTAGCAGGGA  
TAGAACGCCTCTCGCCGGAACCCGCTGACCGCGAAGGCGCGCGGAAGCG  
ATCGAGCGTGTCGGCGCGCACCGGCTCGGCGCCCGTGAAACGCGACCTCCC  
ACGACCGCAGATCGAGCGCCGCTCGCTCCTCCTCCGAGCTCTTCCGGACG  
CACAGGTGCTATGCGAAGTTCGGGCCGCGCTCACCGAGGCGCCGAGCGC  
20 CGAGACGGCGCGGAGCCACCGCATCGGCCTCTGCAGGAACGAGAGCGGGC  
35 ACATGAGCGCGACGCGGATCCGCCGCTAGAGCGCCTGCAAGATCCCGCCG  
ATGAGCCCCATGTCGTGATACGGCGGCAGCCAGATCACCCCGACCGGATC  
CGGGCTCGTCAGGTCGAATCCATGCGCGATGAGCCGCGAGTTGTGCAGCA  
GATTCCCGTGGGTGAGCATCACCCCTTGGGCTCGCCGGTCGAGCCGGAG  
40 GTGTATTGAAGGAACGCGACCGACTCCGGCCGGAGCGCCGCGCCCGGCCC  
25 CTCGATCGGGCCCCGGCGACGGGCCGTGGTTCGCGATCCACCGAGCCGCT  
GCAGCGCGGCGGCCGCGCGCTGGCCGGCAGGGACGCCACGATGCCGGCG  
45 ACGGCCGATGACGTGAGCGCCGCCCTCGGCGCGCGCGTCCGCGACGATGGA  
AGCGACGCGCGGCAGCGTCCGCTCGAGCCGGCCGAGATCCGGCGGATAGG  
30 CGGGCACGGTCCGGACTCCAGCGTAAAGACACCCGAAGAACGCGGTGATG

5

10

15

20

25

30

35

40

45

50

55

TACTCGATCCCCGGCGGATACAGCAGCAGCGCGGGCCCCGGGGGCGAC  
GCCCCGATGCCTGCAAGAGGGCCGCGACGGTTCGCGCGCGCTCGTCAATTT  
CCCCGAGGGTCACCCAGGTGCCCCGGCCTCGACGTCGCCGGACTCAAGA  
AAGCAATAGATTGGGCGGGCGGGCTCAGCTTCGGCCCCGCTGGCGCAAGAG  
5 GTCGATAACGGTGGAAGGGCGGTTCCGTTTCGTTCCGTTCCAATGCAAGAA  
AAGCATCATTCAATTGAACAGACCCCTCCGCCGCGGAGATAGCAGCTTGTC  
CGCTGCGACACAACCGCCGCGCGACGCGCGTGGCACGGCGGGATCCGGGC  
GTTACTCCACCTGCACTTCCCGTCGCGTCACGCTCGCTCCGCCGCGGGTG  
TCGTGAACCACCGCCCACAGCGACACGCGCCCTGGCTCCGAGGGCGGCGT  
10 CCACGTGGTCCCGTTGCCCCCGCGGGAGGCGCCGGTCGTATCGCTCACCA  
GGCGGCGCGCCCCGTCGAACTCGCCACCGTCCGTGTAATAGTCGACCCAG  
ATCGCCTCGCGCGCCGGTGGACCGCCGAGCCCGGGCGGCTTCCTCGTCCAC  
CTCGGCGGCCTTCTCGGGCAGCAGCCTCAATCTCATAGGTCGTGCACT  
CGTCCTCGGCCGGCTCGGTCCGGCCGCGAGCCTTGGGCCTGCTCCTCGGAC  
15 CGAACGCACCGCTTCACGACGGGCAAACCGTCCTCGCCCGGCGCGACCTC  
ATTGCCATCGAGCTTCAGCGTGAAGCCGTCGATGGGCGGGTTCGTGTTCA  
GCCGCTCCTTCTTGAAGACATAGACCTGCGTGTAGCCACGACGAAGCTG  
TCCGGACCGAGCACCGTCCCGTCGTGCGCGACGCACTCCAGCGGAAACCC  
GGCCGTTTCGGGCGCCGAAGCCACGCGTGTGTCGTCGGGCGCACACGGCGA  
20 ACAGCACGTAAGCCGACGAGTACACCGTCCCCGTCTCGGTGGGCCTCGCG  
TCCTTGAGGATCTCCTTGGGCAGCTTCCACCCGAACGAGACCGCATCGGG  
CTCGCCGCTCTTCTCCGGACCGATCTCCTGCTGCGCGAAGGGGACGGTGC  
GCTCCCCGTGCGCGCTGCCGCTGCCGCGTCCGCGCTGCTGCCGCGGTG  
CGGCTGGCACCGCCACCGCCGCGCCGCGCCAGCGCCACCATTGCCGCCTTC  
25 GCCGCCTCCAGCGCCACCACTGCCGCGTCCGCGCGCTGCCACCATCGC  
CACCCTGCGCGCGCGCGCTTCCGCGCTGCCGGCCGGCACCGCCTCC  
CGGATTCGCGACGATTCCCACCGCATAGGTGCCAGCCACTGCGGGATGC  
ACCCGAGGTGCTCGTCCACCCGACCGGCGGATTCACGCAGCCGCCACC  
CACGTGACCTCGACCTTCCGCGGCGCGCCGCCCTCCGCGCCTTTCGCGTC  
30 GGCGTACGTCATCCGGAACGTCACGAGCTCTTCGCGCGCGCGTACGGCT

5

10

15

20

25

30

35

40

45

50

55

TGTCCGCCGTACGGCGAGGACGCGGAGCCCCTTCACCTCGGACGAAGGG  
GCCATGTCGCTCCCGGCGCAGCAAGGGATGCCACGGCCAGGGTCGAGAG  
CGCGGCCAGCAGCGCGCGACGAGCGGGCAGTGCGGTCCGTTTCATCAGAA  
ATCTCCTCGCAGCCCGAGCGTGGGCAGGAAGGGGAGCCCCGTACGTACT  
5 CCGCTTCGTGTAGTTGAAGTTGTAGCTGATGCCCTCCGCAGCCATGTAA  
TTGTAGACGTTCTGGATATCGAGGTAGAGCCCGAGCTGCCACCTCTTGAA  
TTTCACGTCTTGTCGGCGCGGATGTCGAGCTGGTGAAACAGCGGCATCC  
GCTCGCTGTAGTCACCCCCGAGCGGGATCGGCGAATACCTCGCCGAGGAC  
GCGTGGTAGATCGCGTTCACCCGGTTCGGATTGCACCCCTTCTCCTCCGG  
10 ATCGCAGACATAGGGCGTCTGCAGGTTGCCCGACACGAGCCGGAAGCGCG  
CGCCCAGCTCCCAGCCCCGGCCGAGCCGCAGGCTCCCGAGCACCGTCAGC  
ACGTGCGTCTGATCGAACTGGGTGAGGTGCTCCTCCTCGTCGGGGCCGTC  
CTTGCGCACCGACCCGCGAGAGGGTGTACGCCGCCAGCCGAAGAAGCGCT  
25 CGTCCGGCTTGTAATTCAACAAGAGCTCGCCGCCGACCGCGTATCCGGTG  
CCATCGTTGGCATAGTCGTCTTCTCCGGCGAGAAGACGACCAGCCGATC  
GAGCTGCTTGTAAGACCCGTCCAGCGTCACCTCGATCTGCGGCGTGATCT  
30 CCTGCTCCACGCCGAGGCCGTAATGCACGGCGCGGTTGACTTGAGCTCC  
GCATTGCCGAACGGCTCGATGCTCTCCGCGAACTGCGGCGCCTGATAATA  
AAGGCCACGCCCCCTTGGCCGTGTCGCGCGGAAGCCGCTCCGGATGT  
20 CGTAGCGCGCGTTGACCCGCGGGCTCACGTCGAGCGTCTGCGTATCGAGC  
35 GCGTAGTCGACCCGCACCCCGGGGACGATCCGCGCCCGCGGCGAGGGGAC  
GACCTCGAGCTCGGCATACGCCGCGGGCCGCGAGTACGCGCCGTGGAACG  
ACCGATCCTGGAACGGGTACGTCGAGAACGGCTGGTTCGACGGGTGGCCC  
GCGGGCTGCTGCGACGGCGCGCGGATGTTGACCGTGGCGACGCCGCCCGA  
40 GAGGTCGGTGCCGACGTTTCATCGTGAGGTACCGCGCGAACCTGTGCGAGA  
25 GCTCCAGCCGCAGGTCGAGCGAGGTCGAGACGACGTTGAAGGCGAGGGGA  
GAGATCTCGAAGTCGGCGATGTCCCGGCCGAGCGCCATCGACCACAGCAG  
45 CCGATCCCGGCTCCCGATCCGGTTCTCGTAGCTGAGCTGGAAGCGCTGGA  
AGGCGGTGTGCAGCCCGAAATCGCCGTCAGCGCCGGCTCGTCCTCCGGC  
30 GGCTTGTCAGGGTGATCTTGAAGGCGTCGTCCGATCCGTAGAAGCTCGC

5

10

5

15

20

10

25

15

30

20

35

40

25

45

30

50

55

GCGCACGCGCTCGCTCGCGGAGGGGCGGCCCTCGAGGACGAACTGGTAAT  
CATAGTAGACGGGCGCCTGCGTGACGCTGGAGCCCGCCTCCTTGAGCACG  
GGCCCGAGCCACGCGTCGACCCAGCTGCGGCGGCCCGCCGCGATGAACGT  
CCAGTCCTTGAGGAACGGGACGGGGCCCTCGAGGAGCACGCGCCCGTCGA  
TGAGGTCGAGCTGGACCACGCCGTGGTACTTGCCGTCCTGCTTCGGCGAG  
CGGAGCCCGACGTCGACGATGCCGCCCATGGCGCGGCCGTACACGGCGCT  
GAAGTTGCCCGGATAGAAGTCGATCTTCTCGAGCATCTCGGTCGGCACGA  
CCGAGGAGAGGCCGCCGAAGTGGTAGATGATCGGCACCGGGGTGCGATCG  
ACGAACGTGAGCGTGTCTTGGGGCGCGGACCCGCGCACGATGAGCAGCCC  
GAAGCCGCTGCGCGCGACGCCCGGCAGGCTCTGCAGCGACCGCAGCGCGT  
CGCCGCCCGGTGCCGGGGATGCGGTGATCTCGCGGCGCTCGATCGTCCTC  
CGCGTCACCTCGCGCGGCGGGCGCTCGCCCTGCACGGTCACCTCGATGCC  
CGGCGCCTTGCCGTCCTGCGGCGCGGCGAGCGAGATGCGGTAGCGCACCT  
CGATCGCCTCGCCGGCCGCGATCTCCTCCTCGGCGGCGAACGGCTCGAAC  
CCCGCGGCGGCGACCTCGACGCGGTACTTGCCGGGGGGGAGATTCTTGAA  
GCGGAACCTTGCCGCCCTGGTCCGTCTTCGCCTCCTCGCGGCCCGCGTCGG  
GGCGCACGAGGGTGACCGCGATGTCCGGGAGCGGCTCGCCGGTGCCCGCG  
GACAGGACGGTCCCGACCACCGTCTCGACGTCGGCGGGCGGCGCCGGCGC  
GGCCGCATCGGCGGGCTTGGGCGTGAGCGTGAACGCGTACCGGTAGAGGA  
TGCGCGCCCGCGCGGGCGTGCCGTCCGGGCGCCGCGCCGGCGCGAACTCC  
AGGCCGGGCGCGGCCCGCGAGCGCCGCCTCGTTGAAGCCGTGCCCGCCGG  
GCGTCGCGACCTCGGCCTTGGTGACGCGCCCGGTCTTGTCGATGTCGAGC  
TTGAGGATGACGCTGCCCTCGACGCCGGCGCGCTGGGCCTCGATCGGATA  
CGCGGGCGGGGAGTACTTGATCAGCGTCGGCGGGCTGATGGCGGCGGGCG  
CCGGCGGGGGCGCGCCCGGCTGAGGGACGACGACCGCGCCGGCGCCGCGG  
CGGGGGACCGAGGCGCCCGCCGAGTCGCCCTCGGCGGCAGGAGGAGGCTC  
GGGCGGGGGCGCGCCGGCGGGCTGCGCGCGCGCTGCGCTGCCGGTCATCG  
CGACCGCGAGCAGCAGCGCTTCCGAGACGACGAGGCGCATCACGGAGGAC  
GCTGTGGAAGGCATGCGGCCCGCCCTCTCGCATGGCGAGGCCGAGGCGGA  
AAGACGCATCGCGCAGCCAGGACCGTGCTTCACATTGCTTCACACAACGG

5

10

15

20

25

30

35

40

45

50

55

GCGCCGCGCGCTCCCGGGCCGCGCGAGCGCAGGCGGCGCGCGCGCCCG  
CGGGCGGCGGATCGCGAGCGGCGCGGGTGCGATCAGCCGCCGACCTCG  
GCCACGAACCGGCTCACGTCGTCGCTGTCGCCCACGAGCAGCAGCGTGTC  
GCCGTCGCGGATCACGTAGTCCGGTGTGGGCGCCTCGAGCCGCGGCTTGT  
CGCCGGGCGCTTGTTCGTGTGCGGCGGCACACCGAGCACGTTGATGCGG  
TACCGCTGGCGGATCTTCGAGCCGGCCAGCGTCTGCCCCAGCAGCGGCC  
GTGGGCGTTCCAGGGGACCACGCGGTAGTGGCTCGCGAGGTCGAGGAGGT  
CCTGCGCGAGCGGCATGGTGATGTCGGCGCCGACGCGGCGGCCCATCTCG  
GTCTCGAGCTGGATGACGCGGGTCGCGCCCACCGCGCGCAGGATGTCGGC  
CTGGCGATCGGTGGCGGCGCGCGGATGATCTCGCGCACGCCCCATCCGGA  
CGAGGGAGGCCACGCAGAGCACGGACGGCTCGAAGTGCTCGCCGAAGGTC  
ACGATCGCGGTCTCCACGTACTGCGCGCCGATCCCCTCGAGCACCTTGTC  
GACGGTGGCGTCGCCGACGAACGCGGCCGAGGTCTTGTCTTCACGGCGT  
CGACGGCCTCCGGGTGTGTGTCGACCGCGATCACCTCGGCCCGGTTCTTC  
CAGAGGGTCTCGACGACCGACGTGCCGAACCGCCCCGAGCCCCGATGACGA  
GGACGCTCTTCGATTTTCATGGTCTCCGGTCGCGCGCCGCCCCCTCGGGG  
GCGGCGCCGCGAAGGTCTCACGGATGCGCCGGGAGCGCCACGGTTCGGCG  
CTGCCGTCGTCGCGACGGCGCGGCCCGCGCGGGCCGCGCCGCGCGGCGCT  
CAGTAGAGCTCGTCCTGGTGCTGCCACCGCTCCGAGATCCACGCCTTCAG  
GTACGCGATCTCCTCCTCGTACGTCGTGAAGTCGTGCGCCAGCTCCAGC  
CCTCGTAGCTCCGGTACGCCTCGCCCCACCGCGCCTCGTCCCGGCGCGCG  
CTCGCGTCGATGCGCTCCACGTAGCCGTCCACGATCGCGTGGATCTCGGC  
CTCGGCGAGCGCGCCGCGCAGGACCTGATCGTAGCGGGCGCGCAGCGGGT  
CGCCGATCGACGGCTCCTCGAGGAGGCGCTCGAAGAGGAGGTTACGTCG  
CGGTAGTCGACGCGATCCGACGCCGGCTCGCGCTCGGTCTCCACGACTG  
GCCGAAGCTCGCGTTGAAGTCCCACGGCGCGTAGCGGAATACGCCGTCCG  
CGGCCGATCGCGGTAGTGGTAGCTGTTCTTTCCGGCCGAGTCGTTGGCC  
ACGATGAACGTGACGAAGATCCACCAGTCCTCGTAGTCGCGCAGATCGAT  
CCGCGACCCGATCTCGGCGGCGAACGTGGCGTCGTCGGACTCGGCCACGA  
AGCTCACGAGATCTTCCAGATCCGAGAACGCCTCCGGCTCGCCCTCGGCC



5

10

15

20

25

30

35

40

45

50

55

GGCGCCCCTTCCTTCTTCTCGAAGCCGTCGTGCAGCGTGCCTTGGGGTC  
GCCGGACCGGTCGGTCAGCGCGAAGTTCGCGTCGTGGCTGACCGCCTTGT  
AGAGGTTGCCGTCCTGCGGGTAGCCGTGGTCCTCCATCAGGTAGCCGTCG  
ACGTGATCCGCGACGGTGTAGAGCCCCGCGTACTCCCCGTCGAGGTACAG  
5 GACGGCGCTGTAGGTCTTGATCTGGATGTGCTCGGGATCGAGGCGGTTCC  
AGAGGTCATAGGCGAGGCGCTGCCGGACATAGGAGTTGTCGTGGAACGTC  
GTGATGAGCACGACCTTGCGGCGATCGGTGAAGCCGCCCGCCTCGTCGGG  
CTCGTTGAACTTGTGTCCTTGGGGAACCTGAGGGTGTAGCTCCGCTTCG  
GGTACGAGAGCGAGCTCTCGCCGCGGAGCTCCGCCTCCGCGGCGTACGTG  
20 10 TGGCCGCGGTAGATCACCGTGGCCGGGGCGTACTCCTTGTCCTCGGGGAC  
GGGCGAGAGGAAGAGCACCGGCAGGCCGTACTCCTCGGGGTAGCGGGTCG  
GATCGACGACGGGCACGTTCCGACGGATCGGCCAAGGCGTCGGCGACGCCG  
ACCTTGACGCGCCCGACCTCGGACGTCTGCGCGACGCGGATCTCGATGTC  
25 GTAGACGGCGGCCTGATCGAGCCCGGGCGAGAACGTCACCTCGCGCGCGA  
15 TCGGGTCGTACGCGGCGCCCTCGGGGAGCGGGCCGACCTCGAACGCGTCG  
CCGGCGAGCGCGAGGCCGCTCGCGCACGTCACCGGGAACGTCACGGTCTC  
CCCCTCGAGGAGCCAGTGCGGGCCCGCCCGACGGCTGGCAGCGCGAGC  
30 CCTCGGCGCTGGAGCCCCCGCCGCTGGAGCCGGAC

20 Seq ID No 90 (>Contig11)

GGCGACCCACATATCACATAGTAGAATCAGTGTGAGTTAGACAAATGTC  
GAGTGATGAGAAGGACAGAAGTGAGAACTCTGTGATCACTGTAGAACGA  
GAGAGTATGAGCCTGCATACATGATAGCGGACATGAGAACGAGTGTANTA  
TGATGCTACTAAGAGAGTAACAGATCAGAGACTAGAGTAGAGCAATAGAA  
40 25 NTCAGAGATAAGTCAATGACGAGGAGTAGTGATAGAGCTCTTAATAATGG  
CTGAGGTCTGAAGATAGAAGTGATAGAGCGATAGATATACAATCGGTTGA  
AGCAGAGAGTAAGATAAGATCAGACACNGAGTACAGAGAGAGACGAATAG  
45 ATGGCGTGATNTCACAGAGAGGTGCGAGCGTAGCTGACGAGAGCAGAGAC  
CCAGAGTAAGTCACACCTAGATAGTTACGGCGAGAGACAAATGATAGGAA  
30 GGAGTGGACGAGATCAACAGNCCGGAGCACAAGAACGTGAGATGCGACCG

5

10

15

20

25

30

35

40

45

50

55

TGTAATAAACAGGAGACAAGAGCGACTACATAAGAGAGCGAAGCGAATAG  
ATAAGATATAAGCCCAGAGCAAAATAGAAGGAGAGAGAGAGTATTTGTAA  
TAAAGCAACAAGACGGAGAGAGCGAAGCAGCAGGCAACGATTAGAAGAAA  
GACGACAGGAAAGTGAAAGCGAAAGAGAGCAGGTAGAAAGAGAACCAAAA  
5 AAGCACGAAGGAAAAGGAAGCTTCTATGATAGGTGCGGGACAAGGCGTAG  
CTACAGGAGACAGCCGGCATAACGAGGAGCCGGTAAAAGCTAGCCTTTCAG  
AACACATCGGGAGCGCGTAAAGGCGGACCACGCTCGACGGGATCATGTAC  
GCCGACAGCGACGCCTTCAGCCCCGCGCGCACGTCCGGCGCGTCGCCCGC  
CGCGTCGCCGTGAGCACGACGTAGGCGACCAGGCGGGCGTCGCCCGGCG  
10 CGTCCTCGCGCAGGACCACGGCCGCCTGGCCCACGCCGGGCACGCGCCGG  
ATCTGCGCCTCGACGTGCGCGAGCTCGATCCGGTGCCCCCGGAGCTTGAT  
CTGGTGGTCCGAGCGGCCCTGGAACTCGAGCATCCCGTCGGGCAAGAAGC  
GCGCGACGTGCGCCGGTCCGGTACATCCGCCCGCCCGCGGCGCGCGCAC  
GGGTGCGGCAGGAAGCGCTCCGCGGTGAGCCCGGGCTGCCCCACGTAGCC  
15 GCGCGCGAGCGGCGCGCCCGCGATGTAAAGATCGCCGAGCGCGCCGATGG  
CGGGGCGGCGCATCGCGCCGTGAGCACGAACACCTCGGCGTTCGCGACC  
GGCGCGCCGAGGGGGACCCACGTGACCCGCGGGTCGCTCGGCAGGACGCA  
GCCGGTCACCGCGATCGCGGCCTCGCTCGGCCCGTACATGTTGATGAGGT  
CGCCGTGCTGCTTCGCGTAGAAGCGCCGGACGAGATCGAGCGGCACCGCC  
20 TCGCCGCCCCACGAGGACCTTCCGCAGGCTCGCGGGGAACGGCTGCTCGGG  
CCCCCGAGGAACGCCGCGAGCATCGAGGAGACGAAGTACGCGGTGCTCG  
CCCCCTCGTCGCGCACGAGGCGCCGAAGGTACTCGGGATCGCGGTGCCCG  
CCGGCCCGGGCGACGACGATCCGCGCGCCGAACGAGAGGGGCCAGAAGAT  
CTCCCAGACGGAGACGTGAAAGCCGAACGCGGCCTTGAGCAGGACCCGGT  
25 CGTCCGCGGTGAGCGCCCCAGTACCGCTGGATCCACTGCATCTGGTTGACG  
ATGGCGCGGTGGGAGATGAGGCTCCCCCTTCGGCGTGCCCGTCGATCCGGA  
CGTGTAGATGACGTACGCGCCGCTGTCCGGCGGCGGGCTCACGGCGGGCC  
GCGCGTCGGAGCACGCGGCGATCTCGGCGGCCTCGGCGTCGAGGAGCAGC  
GTCGTCCAGCCGCCGGTCGGGAGCTCGTCGGCGATCGCGTCGTGCGTGAC  
30 GAGGAGGCGCGCCCGCGCGTCCCGCATCATGAAGGCGAGGCGCTCGCCGG

5

10

15

20

25

30

35

40

45

50

55

GGTACTCGTGGTTCGAGCGGCAGGTAGGCGCCGCCACCTTGAGCACCGCC  
AGGGTCGCGACGACCATGTCCTCGGAGCGCGGCACGCAGACCCCGACGAT  
CGTGTGAGCCCGACGCCCGGGCGGCGCAGGCAGCTCGCGAGCCGGTTCG  
CGCGCCGCTCGAGCTCGCCGTACGTGAGCGACTTGCCCTCGCTGCGCACC  
5 GCGACGACGTGCGGGTGCTGCTCGGCGCGCTCCTCGAACCACCGGTGCAG  
CGCGCAGGCCGACGGCAGCTCCATCGCGGGGCCGCGCAGCCACGCCTCGA  
15 TCTCGGCGCGCTCGCCGGGGCCGACGTAAGTCCGCTGGGCGACCGGACGC  
TCGGGGTGGCGCGACAGGTCTCGAGCAGCGCCGCGAGGCGCTCGGCGAG  
GCGCTCGGCGTCGCGGCGCGCGCCGCGACGCGTCGTAGCGGAGCTCCA  
20 GCGACGCCGACGGGCCCGCGCCGCGCAGTGCAGCCGGCAGGCGACCTGG  
TCAGACGTGCTCCAGACGTGAGCACGCGGGCCCGCGCGCCGTCGAGCGA  
CAGCGCCGGCGCGGCCTCCGCGGCGGAGAAGCCCCAGCTCATCCGGTGG  
25 CTCACCCCGGGCGCGGCGTCTGCTGCGCGGCGCCCTCGGCCTCGGCGAG  
CGCGAGCCGCGCGCGACGTGCGCGAGCGTGTCCGAGGCCGAGATCTCGA  
15 TCCGCACCGGCAGGAACCGCGCGAACGGCCCCACCGCGCCCGCGAGCGCG  
TCCAGCGACCGCCCGTCGAAGCGGACGGCCACGGTGACCTCGGGCTCGTT  
30 GCGCGCGCTCATCCGCCACAGGAGCGACGCCACAGGGCCAGGAGCACGA  
TCCGCTGCGGGACCTGCCACGACGACGACGACGCGCTCGACCTGCGCCATC  
CCGCCTTGTCAGATCGACCCGCGCGCGCCCCGAGCCGGCGCCGGCGCC  
20 GGCGCCGCGCGGCTGAAGGCGAGGTGGAGCGGGGGCCCGAAATGCGAGC  
35 GGCGCTCGGCCCAGAACCTGCGCCCGTCGCGGCGTCTCCGACTCGAGC  
ATCCCGTTGAGCCACTCGGCGACGTCCGCGTACTGCTGCTCGGGCGGCGC  
GCCCCGCGCCCGCGGTGACGCGCAGAGCTCGCGGACGAGCGGGGCGATCG  
40 ACTCCTCGTCGACGCACCAACGCGGGCGCCGCGAGCACGAGCCGGCGCTCC  
25 TCCGGGCCGACGCGGACCAGGCCGACGCGCAGCCCGTCGTCCGCGCCGCG  
GTCCTCCGAGAGGCGCGCGACGAGCCGCGACATCCGCTCGCCCTGCTCGG  
CTTCGGAGCACCCGACCCAGTCGTCTGCTGACGCCCCACGCGAAGCGCGG  
45 CTCGCCGACCACCTGCGCGGCCTCGCCCGCCCCGCTCGACGAGGCGCGT  
GCGCAAGATCTCGTGCCGCTCGGCCAGCGCGAGCGCCGCCGCGGAGAGCC  
30 GTCCCTCGTCGCACGGGCCGGTCAAGGCGACGACGGCCAGCGTCCGGCAC

5

10

15

20

25

30

35

40

45

50

55

CCGGGCGCCCCGCTCCCGGTCGAGCGCGCGGATCGCCCGCTGCTGCGG  
CGAGAGGCTGAAGCCGGTCATCTCGTGGTCACTCATCCAGGTCGTCCTTC  
GGTGAGGTCTTCGCTTCGCCCCGGCGCGCCCGGGCGCGGGAGGGTCACGGC  
GCGCCGGCGCGCGGAGGTGAGCTTGTCGAGCGCGGCGCCGCGCGCGGCCT  
5 TCCGCTCGAGCTCCCGGCGGGCCGCGCGGCGCGCTCGAGCTCCCCCGG  
AGCTCCGAGACCGGGGTGTCCGGCCGCGCCGTCGCAGTGGCCAGGATCTG  
10 CCGGTAATCGCTAAGGAAATTGTCGACCGTCGCCGCCCGGTACAGCTCGC  
TGCTGTGCTCGACGCCGAAGCGGAACGAGCCGCCGGCCTCGGCGACCGTG  
AGGACGAAGTCGAACGCCGTCGTGGTCGCCTCGCCCTCCAGCGCCTCGAG  
20 CTCGAGCCCCCTCGAGCTTCATCGGGGGGACGTGCACGTTGCGCATGACGA  
10 ACTTCGCGTCGAAGAGGGGCACGTGCCCGACGGCCCCCTTCGGCCGCAGG  
GCCTCGACGAGCCGGTCGAACGGCAGGTCTTGGTGCTCGAACGCCTCGAG  
CGCGACGTCGCGCACGCGGCGGACCAGCGCGCCGAACGTCGGGTCGCCCC  
25 CGCAGTCGGTCCGGAGCACGAGCTGGTTGACGAAGAAGCCGATCATCGGC  
15 TCGGTCTCGACGCGGTTCCGGTTCGCGACGTCGGTGCCACGACGAGGTC  
CTCGAGCCCGGTGCGCTGGTGACGAGGACGAGCTTGACGCGGCGAGCAGGG  
CCATGAAGGGGGAGATCGCCTCCCGCTCGCAGAACGCCTTGATCTGGCGG  
30 GTGAGCTCGGCCCCGGCGTCGAGGCTCCGCCGCGCCCCGCGCCACGTCCT  
TCGCCCCGCGGGCTCGTGGTCGACCGGCACGCGGGCCCGGCGCAGCGCGC  
20 CCGAGAGCTTCGTGTCAGTACCGGAGCTCGCCCTCCAGGACCTCGCCG  
35 GACAGCCACGCCCCGCTGGGCTGCGGCGAAGTCGACGTAAGTCGCGCGGGAG  
CTCCGGCAGCCGGGAAGGCTGGCCCTGCGCGAAGCCGCGGTAGAGCGCGG  
CGAGCTCGCCGACGAAGACGCCGACCGACGACGTCGAACACGACGTGG  
TGACGACGAGCGCGATGACGTGCTCGTCGTGGCGCTTCCGGATGACCCG  
40 CACGCGGAGGAGCGGCCCCGCGGCTCAGGTCGAACGGCGCGAGGCTCTCCT  
25 CGAGGACGAGCGCCGAGACCGCCGCGTCGAGGGCCTCGCCCGCGAGGTGC  
TCGAGGTCGGACATCCGGAACGGCACCCGGGCCTCGGGCGCGACGACCGG  
45 GAACGGCACGCCGTCCCTGGCGCTGAACGTCGTCCGCAGCGCCTCGTGGC  
GCCGCGCGATCTCGAACAGGCTGCGGCGGAGCGCGTCGACGTCGAGCCGG  
30 CCGTCGCGCGCACCACGAACGGGATGTTGTACGCCGGGCTGCCCGGCTC

5

10

15

20

25

30

35

40

45

50

55

GAGCTGATCGACGAACCACAGCCGGTGCTGCGCGAACGACAGCGGGAGCG  
GGCCGTGCGGGGGGATCCGCGCGATCGGGGGGAACGCGCCGCGCGCC  
TCGCCGCGGCGCGCGGCGTCGACCTGGGCCGCGAGCGCCGCGACCGTCGG  
CCCCTGGAAGAGCGCCCGCAGCGGGAGCTCGACGCCGAGCTGCGCGCGGA  
5 TCCGGGACATCACCTGGGTGCGGACGAGCGAGTCGCCGTGCAGGCCGAAG  
AAGTCGTGTCGACGCCGATCTCGTGACGCCGAGGAGGGCGCTCCCAGA  
TCGCGGGCGATCGCGCGCTCGGACTCGGTGACGGCGCGGCGAACGCGGC  
CCCGGCGTGCGCGCGGAGACCGCCGACGTGCGGAGGGCGTTCGGCGCGG  
GCGCGAGGGGCGCGTCCGCCGGCGCCGGCTCGGCGGCGGGGCGCCGCGG  
10 CGCGGCGCGATCCAGTGCCGCGCCCGCTCGAACGGGTACGTGCGCAAGCG  
GACGAGCGCGCCGGGGACGCCCCCGCGGACGGGCCGTCCAGTCGACGGC  
GTGGCCCGCCTCCCAGAGCTGGCCGAGGGCCTCGGCCAGGCTCGCGGGCT  
CGGACGCGGCGTGGGTGACCCGAGGCTCGCGATCGCGGCGCCCGCGCG  
15 CCGGCCAGCGTCTGCCGCACCAGCGTGGTCAGCCCGCGGCCGGGGCCGAC  
CTCGAGGAACAGGGCGTGCCCGACGCGAAGAGCGCCTCGACGCCGTGCG  
TGAAGCGGACCGGCTGGCGGAGGTGCCGCGCCAGTAGGCCGGATCGGTC  
20 GCCTCGGCGTCGGTGAGGAGGGCGCCGGTGACGTTGAGACCACGGGGAT  
CTCCGGCGGGGAGAGCCGCGCGCGCCGCACGCTCTCGAGGAACGGGGCCA  
CCGCGCCGTGATGAGCGCGCAGTGGAACGCGTGGGACGTCTGCAGCGGC  
25 CCGGCGAACACCTCGCGCGCTCGAGGCGCGCGGCGAGATCGCGGATCGC  
GCTCGCCGGGCCCCGCGACAACCGTGAGCTTCGGGCTGTTGACCGCGGCGA  
TCTCCAGGCCGGCCTCGAGGAGGCCCTCGACGTCCGCGGCCGGCAGGCCG  
ACGGCCAGCATGCTCCCGGCCGGCGCCGCTGCATGAAGCGCCCCCGATC  
30 GATGACCAGGGACATCGCGTCCTCGAGCGTGAACACGCCCGCGACGCAGG  
CCGCCACGAGCTCGCCGAGGCTGTGGCCGATCATCGCCGCGGGCTCGATC  
CCCCAGCTCATCCAGAGCCTGGCGAGCGCGAGCTCGACGGCGAAGAGCGC  
GGGCTGCGCCAGCGCGGTGCCGAGCAGCGTGCGCCCGTCGCCCTCGCCCT  
45 CGCGGAAGACGACCTCGCCGAGATCGAGGCCGCGCGCCCGCGCCCGCC  
GCGCACGCGTTGAAGGCGCTCCGGAACGCCGCTCCTGCGCGTAGAGCGC  
30 GCGGGCCATCCCGACGGCCTGCGCGCCCTGGCCCGGGAACGCGAAGACGG

5

10

15

20

25

30

35

40

45

50

55

GCGCGGCTCATCGGGGCGCGGAGCGCGCTCGCCCCCTCGCGGGCGAGCC  
CCTGGATCGCCTCGGCGCGCGTCCGGGCGACGACCGCCCGCGGTACGGG  
TGCTCCGCGCGCCCGGTCTGGAGGGTGAACGCGACGTCGTCGAGCGGGAC  
GTCGGTCGCCTCGAGGTGCGCGGCGAGCTGCGCGCAGGCCGTCGACAGCG  
5 CCTCCGGCGTGCGCGCCGAGAGCGTCAGCACGTGATCGCGCTCCGGGGCC  
GGGGCGCGGGCGGCAGCGGGGGCGGCTCCTCGAGCACGACGTGCGCGTT  
15 CGTCCCGCCGATCCCGAACGAGCTCACGCCCGCGCGGCGCGGGCGGAGCT  
CGCGGGGCCAGGGCGCCGCTCCCGCGGGACGAAGAACGGGCTCGCCGCG  
AGGTCGAGCTTGGGGTTCGGCGCCTCGAAATGGACGCAGGGCGGGATCTC  
20 GCCGCTCCGCACGACGTGCGCCGCTTGATGAGGCCCGCGACGCCCGCCG  
CGGCGTCGAGGTGGCCGATGTTGCGCTTGATCGAGCCGAGCGCGCAGTAC  
GCCTTCCTCGGGGTCTTGCGGCGGAAGGCCTGCGTGAGCGCCTCGACCTC  
25 GATCGGATCGCCGATCGCGGTGCGGTCGCGGTGCCGTGGGCCTCGACGTAGCCGA  
TCGAGCCGGGATCGACGCCGGCGACCGACTGCGCCTCGGAGATCGCCGCC  
15 GCCTGGCCGTCGACGCTGGGCGCCATGAAGCCGACCTTGCGCCCGCCGTC  
GTTGTTGACGGCGGAGCCCCGATCACCGCGTGACCGTGTTTCGGTCGC  
GGAGGGCGTCCGCGAGGCGCTTCAGCGCGACGATGCCGACGCCGCTGCCG  
30 CCCACGGTCCCCCTCGGCGCGCGCGTCGAACGGCCGGCAGCGGCCGTCCGG  
GGAGCAGATGCTGCCGGGCACGTACGGATACCCGCGCTTCTGCGGGATGC  
20 CGATGGAGACGCCCGCCCGCCAGCGCGAGATCGCACTGGCCGCCGAGGAGG  
CTCTCGCACGCCATGTGGACGGCCACCAGCGACGTCGAGCACGCGGTCTG  
CACGACCACGCTCGGCCCCGTGGAGGTGAGTTTGTACGAGACCCGCGTCG  
CGAGGTAATCCTTCTCGCTCGCCAGCATGAGCGCGTGCGGATCGACGGTG  
40 GCCGCGAGATCCGGGTGCGAGAGGAGCTGGAGGAGGTACGTGTTGGAGCC  
25 GCACCCCCCGAAGACGCCGATCGCGCCCGGGAACCGGGCCGGATCGCAGC  
CGGCGTCCTCCAGGGCGGCGACCGCGCACTCCAGGAAGAGGCGCTGCTGC  
GGGTCCATGAGCTGCGCCTCGCGCGGCGAGTACCCGAAATAGGACGCGTC  
45 GAAGCGGTCGATGTCGTCGAGCAGGCCGCCCGCGCAGACGAGGGCGCCC  
CGGGGGCCGCGCTCGCGCCGACCGGCGGCTCCTCGCGCTCGCTCTCCGGG  
30 AAGCGCGGATCGACTCGACGCCGCGCCGCACGTTCTCCAGAGGGCGTC

5

10

15

20

25

30

35

40

45

50

55

GACGCTCGGGGCGCCGGGGAAGCGGCCCGCCATGCCGACGATCGCGATGT  
CGCTCCCCCGTCTCTCGGTCTCGATCGGCTCTGACATGGCTATCCTCGCC  
CCCCGGCGGCGTCTCGCGGTCTCGGGCGCGCCTCGGCGCGCTGCGCCCCGACG  
TCGGCCGGCTCGGCCTTGACCGTCTCGCGGTCTGAGCCGCTGCGCCAGTTG  
5 CTCGATGGTCGGGTACTGGAACAGGTCTGGTCAGCGACACGGCCTGCGCCC  
CGGCGCCCTCGTCTGGGCGCGCGCGCCGCGATGCGCTCGGCGAGCAGGCGC  
15 TGCGCGCGCACGAGGAGCAGCGAGGTGAAGCCGAGCTCGAAGAGGTTGTC  
GGTCACGCCGACGGCCTCGACCTGCAAGACCTCCGCGAGCACCGAGGCGA  
TGAGCCGCTCGGTCTCGGTCCGCGGGGCGACGGCCGCGGCGCGCGCGCG  
20 ACCGCGGCGGGATCCGGCAGGGCGGCGCGGTCCACCTTGCCGTTCTCGGCT  
10 CAGCGGCAGCGCCGGGAGGACGACGACCTCCGCGGGGATCATGTACTCCG  
GCAGCTTCTTCGGACGAAGTCGCGGAGCGCGGCGCCATCGCCGTCTGGCG  
CCGACGACGTACGCGACCAGGCGCTTCTCGCCCCGACGGATCGGTCTTCGC  
25 CGCCACGACCGCCTGCTCGACCGAGGGGTGCTGCGCGAGGGCGGCCTCGA  
15 TCTCGCCGAGCTCGATGCGGAAGCCGCGGATCTTCACCTGATGGTCGGTG  
CGCCCGAGCAGCTCGATGGTCCCGTCTGGCGAAGTAGCGGCCAGGTCTGCC  
30 TGTCTGTACAGCCGCTCGCCGGTCTGTGGGGTGCTTCAGGAACCGCTCCC  
GGGTCCGCGCCTCGTCTCGCGCCAGTATCCGAGCGCGACGCCGATCCCGCCG  
ATGTGGATCTCGCCGGGGACCCCGATCGGACACGGCTCCAGCCCCTCGTC  
20 GAGCACGTAGGTGTGCTGGTTCGCGAGCGGGCGGCCGTAGGGGATGCTGC  
35 GCCACGCCGGGTCTGACGTCCGCGATCGGGTGGGCGATCGACCAGATCGAC  
GCCTCGGTCTCGCGCCGCCGAGGCTCACGACGCGGGGCGCGCGGCAGGCCGC  
GCGGATGCGATCGGGGAGCTTCAGCGGGATCCAGTCGCCGCTCATCATGA  
CGAGGCGGAGCGACGACAGCGCCGGTCTCGCCCGCGCCGGGGACGCGTCC  
40 ATGAGCATCTCCATCAGCGCCGGGACCGAGTTCACACGGTCACCCGCTC  
25 GCGCTCCACGAGCTCGCGCCAGTGCCCCGGATCCGAGGCGCGGGTACGGT  
CGGGGATCACGACGGCGCCTCCGGCGGCGAGCGTCCCGAACACGTCGTAG  
45 ACCGACAGGTCGAAGCTCAGCGACGAGAGCGCGAGCACCCGGTCTCTCCGG  
GCCGACGTCGAAGCGGCGGTTGATGTGAGGACCGTGTTCACCGCGCCGC  
30 GGTGGTCGATCATCACGCCCTTGGGCAGCCCCGTGGACCCGGACGTGTAG

5

10

15

20

25

30

35

40

45

50

55

ATCACGTAGGCCAGGTCGTCCGTGCTTCCGCCGGGCGGCCGGCGCGCGAC  
GGGCTGCTCGCGCCACCGCTCGTCCGCGTCGACGGCGAGGCGCTCGATGC  
CCGCGGGCCAGGCGATCGTCCCGTCGACCGCCGACTGCGTGAGGACGAGG  
CGGACCTCGGCGTGCTCCAGGAGGTGCCTGAGGCGCTCCTCGGGGAGGCG  
5 AGGGTCCAGGGGCAGGTAGGCGGCGCCGGCGCGCAGCACGCCGAGCACGG  
CGGCCACCTGCTCCCAGCCCTTCTCCATGACCACGGCGACGAGCGCGTTC  
15 GCGGTGCTCCGGAGCGCGAGGCCGCGCGGCGATCGCCTCGGCGCGCCG  
GGCGAGCTCCCCGTAGGTGAGGCGCCGCTCGGCGTCGACGACCGCGCACG  
CGTCGGGCTGCTCGACGGCGCGCTCAAAGAACGGCTCCTCCAGCCGGAGG  
20 TGATCCGGGGTTGCGACCGCGGTGTGCTTCCACGCGACGAGGGCGCGCTC  
10 GCGGTCTTCCGGCGCGACGGAGAGCGCGCGGACGCGCTGCGCGGGGTCT  
GAGTGGCGCGGAGAGCACGCTCTGCATCGTGGCGAGCATCCGGTCGATG  
GTCGCCGCGTCGAAGAGGTGACGTTGTACTGGAGCGAGATCACGTCGCG  
25 CCGCCGCGCGGCTCGACGCTGAAGCGCAGGTGGAAGCGCGTGGCCTCGA  
15 CCGGGAGATCGAGCGGCTCGATCCGCACCTCGCCGAGCTCGAGCGCCCTCG  
GTTGGGGCGTTCTGCACGACGAGCATGACCTGGAACAGCGGCGAGCGGCT  
CAGGTCGCGGCGGGGTTGACCGCCTCGACCACCTTCTCGAACGGGGCGT  
30 CCTGGTGCTCGAACGCCTCGAGCGCGACCTTCCGCGCCCGCGAGAGGAGC  
TCCTCGAAGGTCGGGTGCGCCGCGAGGTGAGGCGCATGACGATCGTGTT  
20 CACGAAGAAGCCGACGAGGGGCTCGAGCTCGGGGCGAGGCCGGTTGGCGA  
35 CCGCGGTCCCGATGGCGAGGTGCTCCTGGCCCGAGCTGCGCCGGAGGAGC  
ACGCCGAGGGCGGCGAGCAGGACCATGAAGCGGGTGGCGCCGCGGCTCCG  
GGCGAGCTCGTCGAGCTGCGCCACGAGGCGCGCTCGAGCGGGAGGACCC  
GCTCCGCGCCCGGAACGTCTGGACGGGCGGCCGCGGTGATCGGTCTGG  
40 AGCTCCAGGACCGGCAGCCCGCGGAGGGTCGCTGTCCAGTGAGCGAGCTT  
25 GTCGGCGAGCCGCTTCCCCGCGAGGTGGCGGCGCTGCCACACCGCGAAAT  
CGACGTACTGGAGCGGCAGCTCGGGCATGTCCGCGGGCCCGCCGCCCCGC  
45 GCGCGCCGGTAGAGCTCCGCGAGATCGCGGACGAGGGGTTGGAAGGACCA  
GGCGTCCGTGACGATGTGGTGCGTGACAGGACCAGGACGCAGACGTCGT  
30 GGTGAGGCGGAACAGCCTGGCGCGGAACACGGGCCCGCGCGGAGGTGCG



200

5

10

15

20

25

30

35

40

45

50

55

AACCCCGTGGCCTGCTCGCGCGACGCCAGGCGCGCGCCGCGGCCTCCGC  
CTCGTCCGGGGGCGTGCCGCGGAGGTCCACCACCTCCGCGGGGGCCGCCT  
CGGGCTCGCAGATCTTCTGCGCCGGCGTGGGGCTCGCGACGAACACCGTG  
CGCAGGCTCCAGTGCCGCCGGACGAGCGCGGCGAGCGCGGAGGACAGCGC  
5 GTCGACGTGACGAGGTTCCGCAGGCGGACCGCCTGCACCACGTTGTAGG  
CGGTCCCGCCGGGGAGCAGCTGCTCGAGGACCCACAGGCGCTCTTGCTCG  
15 TACGAGAGCGGATACGGCTCGTCCGCCGGCGCGCGGCCCCAGCGAGGGCGC  
GATCTCGCTGGCGGGCACCGTCGCTGCGGCGGCGGTGGTCGAGGCGGCGC  
CGGAGGAGAGGCGATCGGCGAGCTGGTGGAGGGTTGGGTGCTCGAAGAGC  
20 GTGCGGAGGGTGGTGCGGATGCCGAGGGAGGACTCGATGCGTCCGAGGAC  
10 CTGCATGGCGAGCAGGGAGTGGCCGCCGAGGTGAAGAAGCTGTCGTGTC  
GTCCGACGCGGTGAGGTGGAGGACGGATTGCCAGATGTGCGCGAGCTCC  
CGCTCGAGCTCGCCGAAGGGGGCTCGTAGTCGGCGTGCGCGCGGGTGG  
25 CGCAGGGAGGAGCTTCTTGTCGACCTTGCCCGAGAGGGACATGGGCAAGG  
CGGGGAGCAGGACGAAGTGGGCGGGCACCAAGGCGTCGGGCACCAGGCGG  
15 GCCATGCCCTCGCGCAGGTCGCGCTCGGAGGGCGGGTCCGCGCCCGGCAC  
GACATAGGCAATCAGGCGCGCGGCGCTGCCTTGGCCGTGGAGGACGACGA  
30 CGCCCTCGCGGACGGCGGGCAAGCGTCGCAGGGCGGATTGACCTCGCCG  
AGCTCGACACGGCGACCGCGGAGCTTGACCTGCTCGTCGCGGCGTCCGGC  
20 GAAGGCGAGCTGTCCGTGCGGGCGCCAGCGCACCAAGGTCCCGGTGCGGT  
35 AGAGGCGTGCGCCGGGCTGGCCGAAGGGATCGGGCAGGAAGCGCTCTGCG  
GTCAGGTCCGTGCGTGTGTAGCCCTGGGCGAGGCACGCTCCGCCGATGTA  
CAGCTCGCCGAGGACGCCGGGCGGGACGGGCTGCATGTGCGGGTCGAGGA  
CGTAGACGAGGGCGCTGTGATGGGTGCGCCGAGCGGGGGCTCGTCGCCG  
40 AGGTGCGGCGACCTCGGCGACGGTGGTGATGACGGTGGCCTCGGTGGGGCC  
25 GTACATGTTGAAGAGGCGGAAAGGGAGCGGTGCGCCGAGCGGATGGAGCT  
TGTCGCCCGCCGACGGTCATCGCGCGCAGGGCGATGCCGGTCCAGTCTTGC  
45 TCGAAGCACGCCTCGGCCAGGGGCGTGGGCATGAACGAGAGCGTGGCCCG  
CTGAGCGACAAGCCAGGAGACGAGCGCTGTGGGAGAGCGGAGCGCGTCGT  
30 CGTCGGCGAGGAGGAGTGCAGCGCCGCAGGCGAGCGGCGTCCAGATCTCG

5

10

15

20

25

30

35

40

45

50

55

5 TAGACGGAGGCGTCGAAGCCGCTGGAGGCCAGCTGAGTCCAGCGATCGCG  
GGGTGAGAGCGCGAGCAGGTGCTGGAAGAAGGAGACGAGCCTTGAAAGGC  
10 TCGCATGGCGCACACAGACGCCCTTCGGCGTGCCGGAAGAGCCGGAGGTG  
AAGAGGACATAGGCCAGGTCGTCTGGGCCTGGAGACGAGAGGAATGTGGGT  
5 GCTGGGCGCGCACGCCCCGTCTTGACGAGGTGGACGGGGCAGGGGGCGG  
CGGTGAGCTTGTGGCTGGCCTGGCTGCTGGTGAGCACGAGCGCGGCGCGG  
15 CAGTCGGCGAGCATCTCGGCCAGGCGCGCCGGGGGGTTGGCGGGGTCGAG  
CGAGGCATAGGCGGCGCCTGCCTTGAGGACGGCGAGCTGGGCGGCGACCA  
TGCGGGGCGAGCGCTCGATGCAGACGCCGACGACGCTGCCGGGGCCGACG  
20 CCGCGGTTCGCGCAGCCACAGGGCGAGCTCGGTGGACCAGGTGCTGAGCTC  
10 TGCGTAGGTGAAGCGCTGGTGTCCGAACTCGAGCGCCGTGGCGTCCGGCT  
GTCGAGCGGCGTGGGCCTCGAAGAGCGCATGGACGCAGGCGGGGGCCGGG  
GCGGAGGCGGCCTGTCGTGCGGCGGCAGCGCCGCTCCAGTCGTGAGGAG  
25 CAATGCGCGCTCGGCGTCCGAGAGCATCCGGAGCTCGGAGAGCGGTGAC  
15 CGGGGTGCTCGACGGCGCTTTCGAGCAGGAGCACGAAGTGGCGCGCCATC  
CGCTCGATGGTGGCGGGGTCGAAGAGCTGCTGGTCTGACTCGAAGCGCAG  
GGCGATGCCGAGTCGAGCTCTGCCGCGAACAAGGCGAGATCGAACTCGG  
30 CCGCTGCCTGCTCGTCGGCGAGCGTGGTGAGCTCGAGCTCTCCCTGCGCG  
ATCCGCACGTCTTCCGCGCCGGTGGTGAGCGCGGCGAGGCGGGGATCCAG  
20 CGACGGCAGAGCGCCCTGGAAGGCCAAGGCGACGTCGAAGAGCGCGCCGC  
35 CTCGCCGGGCGCGCCCCGGGGCTCTGCGAGCAGGTGCTGGAGGGCGCTG  
TCGCCGTGGGCCAGCCCGTCGAGGAACGCGTCGCGCACGCGGGCGACGAG  
CGCGTCGAAGGACGCGGCCCCGCGCAGCGCCACGCGCACGGGGAGCATCT  
GGACGAAATAGCCGAAAGCCCCGAGTGCTCTCGTCGTCTGTTCCGCCCCGCC  
40 GAGGGGACGCCCACGACAAGGTCGTTCTGCCCCGCTCGCGCGATGGAGCAA  
25 GACGGTGAGCGCCGACAGCAGGACCGAGAAGAGCGTGGTCCCGCGCTCGC  
GCGCGAGGCGCGCCAGCGCTCCGGTCAGGGGCTTTGGCAGCGTGATCGCG  
45 TGAGCGCGACCGCGGCGAGGGCTCGCGTCGTGGCGGGCCCGGTCCGCGGG  
AAGGTCGATGGCGGTCTGTCGCGCCGTCGAGCGCCTTGCGCCAGTATTCTG  
30 CTCCGCCGGCGCCTCCCGCGGCGAGGGACAGCTCACGCCGGCGGCGAAG

5

10

15

20

25

30

35

40

45

50

55

AAGCTCGACGGCGGCGGCAGCTGCGGGGGCCGGCCCGCGCGCAGCGCCGA  
GTACAGCTCCCCAGCTCGCGAACGAGCAGCGCGAACGACCAGTAGTCGA  
CCACGAGGTGGTGAACGACCACCGTGAGCAGCGGCGGCTGCCCTCTCCG  
CGCCGCCAGACATGCACCCGGAGCAGCGGTCCGCGCTCCAGGTCGAACGC  
5 GCGGCGGCGCACCTCGTCCGCGCGGGCGACGATCTCGCGCTCGTCCAGCG  
CCATCGCCGGCTCTTCGGCCCATTCAGGGCGACATGGCGGTGGACCTGC  
15 TGCAGCGGATGGCCGTGCGCGCTGAGGAACGTCGTGCGGAGCGCCTCGTG  
CCGCTCGACGAGGCCCTCGAACGCGCGGCGCAGCGCGGCCACGTCGACGC  
CGGCACCGAGCCGGACCGTCCTGCCCAGGTTGTAGAGCGCGCCGTGCGCC  
20 GACTTCTGGCACTCCAGCCACATCGCCCGCTGCCCTCGGTGAGCGCAA  
10 CGGCTCTTCCGGCGTCCACACCCGCGGCGACGAGGGGCCGTCCGGACGGG  
TGAGCCCTCGCTCCAGGGCCGTGCTGCGGCGGCGGAGGTCGAGGCGGCG  
CCGGAGGAGAGATGACTGGCGAGCTGCGCGAGGGTTGGGTGCTCGAAGAG  
25 CGTGCGGAGGGTGGTGCGGATGCCGAGGGAGGACTCGATGCGTCCGAGGA  
15 CCTGCATGGCGAGCAGGGAGTGGCCGCCGAGGTCGAAGAAGCTGTCTGT  
CGTCCGACGCGGTGAGGTGGAGGACGGATTGCCAGATGTGGGCGAGCTC  
GAGCTCGAGCTCGCCCGAGGGGGGCTCGTAGTCGGCGTGCGCGGCGGGGG  
GCGCAGGGAGGAGCTTCTTGTCGACCTTGCCCGAGAGGGACATGGGCAAG  
GCGGGGAGCAGGACGAAGTGGGCGGGCACCAGGGCGTCGGGCACCAGGCG  
20 GGCCATGCCTTCGCGCAGGTGCGGCTCGGAGGGCGGGTGGGCGTCTGGCA  
35 CGACATGGGCAATCAGGTGCGCGGCGCTGCCTTGGCCGTGGAGGACGACG  
ATGCCCTCGCGGACGCCGGGCAAGCGTCGCAGGACGGATTTCGACCTCGCC  
GAGCTCGACGCGGCGACCGCGGAGCTTGACCTGCTCGTCGCGGCGCCCCG  
CGAAGGCGAGCTGTCCGTGCGGGCGCCAGCGCACCAAGGTCGCCGGTGCGG  
40 TAGAGGCGTGCGCCGGGCTGGCCGAAGGGATCGGGCAGGAAGCGCTCTGC  
25 GGTGAGTCCGTGCGTGTGTAGCCCTGGGCGAGGCACGCTCCGCCGATGT  
ACAGCTCGCCGAGGGCGCCGGGCGGGACGGGCTGCATGTGCGGGTCGAGG  
45 ACGTAGACGAGGGCGCTGTGACGCGGTGCGCCGAGCGGGGGCTCGGCGCC  
GAGGTCGGCGATCTCGGCGACCGTGGTGATGACGGTGGCCTCGGTGGGCC  
30 CGTACATGTTGAAGAGGCGGAAAGGGAGCGGGCGCCGGAGCGGATGGAGC

5

10

15

20

25

30

35

40

45

50

55

TTGTCGCCGCCGACGGTCATCGCGCGCAGGGCGGAGCCGGTCCAGTCTTG  
CTCGAAGCACGCCTCGGCCAGGGGCGTGGGCATGAATGAGAGTGTGGCCC  
GCTGAGCGACAAGCCATGAGACGAGCGCCGTGGGAGAGCGGAGCGCGTCG  
TCGTGCGCGAGGAGGAGGGCAGCGCCGCGAGGCGAGCGGCGTCCAGATCTC  
5 GTAGACGGAGGCGTCGAAGCCGCTGGAGGCAACCTGAGTCCAGCGGTTCGC  
TGGGCGAGAGATCGAGTCGGAGGTGGAGGAAGGAGACGAGCCTTGAAAGG  
CTCGCATGGCGCACACAGACGCCCTTGGGGGTGCCGGTGGAGCCGGAGGT  
GAAGAGGACATAGGCCAGGTTCGTGGGCGCTGGAGACGAGAGGAATGTGGG  
TGCTGGGCGCGCACGCCCCGTCCTGGACGAGGTGGACGGGGCAGGGGGCG  
10 GCGGTGAGCTTGTGGCTGGCCTGGCTGCTGGTGAGCGCGAGCGAGGCGCG  
GCAGTCGGCGAGCATCTCGGCCAGGCGTGCCGGGGGGTTGGCGGGGTCGA  
GCGAGGCATAGGCGGCGCCTGCCTTGAGGACGGCGAGCTGGGCGGCGACC  
ATCGGGGGCGAGCGCTCGATGCAGACGCCGACGACGCTGCCGGGGCCGAC  
25 GCCGCGGTTCGCGCAGCCACAGGGCGAGCTCGGTGGACCAGGTGCTGAGCT  
15 GTGCGTAGGTGAAGCGCTGGTGGCCGAACCTCGAGCGCGGTGGCGTCCGGC  
TGTCGAGCGGCGTGGGCCTCGAACAGCGCGTGGACGCAGGCGGGGGCCGG  
GGCGGAGGCGGCCTGTCGTGCGGCGGCAGCGCCGCTCCAGTCGTGAGGA  
GCAATGCGCGCTCGGCGTCGGAGAGCATCCGGAGCTCGGAGAGCGGTCTGA  
CCGGGGTGCTCGACGGCGCTTTCGAGCAGGACCACGAAGTGGCGCGCCAT  
20 CCGCTCGATGGTGGCGGGGTCTGAAGAGCTGCTGGTCTACTCGAAGCGCA  
GGGCGATGCCGGCGTCGAGCTCTGCGGCGAACAAGGCGAGATCGAACTCG  
GCCGCTGCCTGCTCGTCGGCGAGCGTGGTGAGCTCGAGCTCTCCCTGCGC  
GATCCGCACGTCCCCACGCCGATCGCGAGGGCTGACAGGCGTGCATCCA  
40 GCGATGGCGGGGTGCTCTGGAAGGCGAAGGCGACGTCGAACAGCGCGTCT  
25 CGCTGCGCCTCGCCCTGCGCTCGCGCGAGCAGGTGCCGGAGGGCGCTGTC  
GCCGTGGGCCAGCGCGTCGAGGAACGCATCTCGCACGCGGGCGACGAGCG  
CGTCGAAGGACGCGGCCCCGCGCAGCGCCACGCGCACGGGGAGCATCTGG  
ACGAAGTAGCCAAAGGCCCTGGCGCTCTCGTCGTCGTGCCGCCCCGCCGA  
45 GGGGACGCCCACGACAAGGTGCTCTGTCCGCTGGCGCGATGGAGCAAGA  
30 CGGTGAGCGCCGACAGCAGGACCGAGAAGAGCGTGGTCCCGCGCTCGCGC

5

10

15

20

25

30

35

40

45

50

55

5 GCTCGACGGCGGGCGGCAGCTGCGGGGGCCGGCCCGCGCGCAGCGCCGAGT  
10 GGTTCGATGGCGGTTCGTTCGCGCCGTCGAGCGCCTTGCGCCAGTATTCTGCT  
CCGCCGGCCGCTCCCGCGGGCAGGGACAGCTCACGCCGGCGGCGAAGAA  
5 GCTCGACGGCGGGCGGCAGCTGCGGGGGCCGGCCCGCGCGCAGCGCCGAGT  
15 ACAGCTCCCCCAGCTCGCGAACGATCAGTGCGAACGACCAGTATTGACC  
ACTACGTGGTGATCCACCACCGTCAGCACTCTGCGTTTAGTCCTTTCTCC  
TGTTTCGGCCTATTAATTGCTACTATGGATCCACACTGCTCCGCCTCTTGT  
ATCTCCCTTATCTGCACTTGCTGCGCTTNACCTTTACGTTCCCTCTCCCTG  
20 CTCTATACTATTTCTTCCCCCGCTTCTCGTCCTATTCTGCATTTGTCATA  
TCGTATCTTCATATAACCTTTCTTTTCGCTATCCTTACTGCTTCTCGACCTT  
ATGTGCGTCTGTCTTCCCTTTCTNTATTATTCTCTGTCTCACCGCTCTN  
TGCTCTGTGCTCCTATCACTAAATTATGTCTCTATCACTGCTACTATCT  
25 GAAGCTGATCTTCGAGATCTCGCTNGGTGTCACTCTTTATCTCATAGNCG  
15 CCTCTGTCTTCTTGTCTCCTTAAGNCTGATTTTCTCGCTCTATTCGTGAC  
TACTCTGCTGTCTCTCACATACGTGTTCTTGAATCGTATTTCGCGTTCTCG  
CTACTGTGATATCCATTGCCGACCTCTACTGCTCNTCTNTATGCTATACT  
TCTTAGTCTCTTACTACGTTNGTCTGATATNTTGCTGACGACGTCATGTC  
ACGCTCGCAACTCTTCANTTCTATCGTATACGCTGATCATTTTTCTGT  
20 GAGGCTGATGTACTATACGTAATTACCTGTATACGTCGTCTATCTACTCT  
35 CGTGTCTTCACTCTTTCTACTCC

Seq ID No 91 (&gt;Contig12)

40 CCCCCCGCCGTCCGCCGGTACGTTCGCGGACCGCCGCCCGAGCAGCTCCC  
25 CGCGCTCGCGCCGGAGGAGCGGGAGGCCGCGGCGCGCCGCTGTGCGCCC  
TCGGCGCGGCGCCGCCGAGGTCCGGCGCCGCGGGCTGACGCGGGCGCCG  
CTCTCGTACGGGCAGAGCCGCATCTACTTCCTCGAGCAGCTCTCGCCCGG  
45 CAAGCCGCTCTTCAACGTCCCGGGCGCGGTCCGGCTCCGGGGCCCGGTTCG  
ACGTCGCCCCGCTCTCGGCAGCGTTCGGCGAGATCGTGCGGCGCCACGAC  
30 GCCCTCCGCACGTTCGATCGCCAACGTTCGACGGCGAGCTCCTGCAGATCGC

5

10

15

20

25

30

35

40

45

50

55

GCAGCCGCACGCGGGCTTCGCGCTCGACGTGGTGACCTCGACGCCCCGAGG  
AGGCGGCCGAGCTCGACCGGCGGCTGCGCGCCGAGGCGTGCGGGCCCTTC  
GCGATCGGCGCGCCGCGCTCCTGCGCGCCACGCTGTTCCGCCTCGCGGA  
GGACGAGCACGTGCTCCTCGTCACGATGCACCACGTGGTGTGCGACGACT  
5 GGTGCTCGGCGTGATCCTCCGCGAGCTCCTCGCGCTGTACGCGGGCCGC  
TCGCTCCCGCCGCGCGGCTCCAGGTACGCGACTTCGCGGCGTGCGAGCG  
15 CGAGATGGTCGAGTCGGGGGCGCTCGACGGCCAGCGCGGTACTGGCGAG  
AGCGCCTCCGGGGGCTGTCCCGGGCGAGCATCTCGGCCGGCGGGGGCG  
GAGGCGCCGAGCCACGACCCGTCCGGCGCCATCGAGGAGATCGCGCTCTC  
GCCGACAAGGCGGCGGCGCTCGAGGCGCTCGCGCGGCGGGAGGGAGCGA  
20 CCCTGTTTCATGGTGCTCCTCGCGCTCCTCGACCTCGTGATCCATGCGCGG  
TCCGGCGCACTGGACATCGCCGTGSGGACGCCCATCGCCAACCGGAACCG  
CCCGAGCTCGAGGACGTGGTCGGCCTCTTGACGAACACGCTCGTGATCC  
CGGTCGATCTCGCGCGCGCCGGGGCGTTCCGCGACGTGCTCGCGCGGGCG  
15 CGCGTCCAGGCGCTCGACGCCTTCGCGAACCAGGACATCCCGTTTCGATGT  
CGTCACCCAGGATCTGAAGCAGGAGCGCGACCACGCGCAGCACCCGCTCT  
TCCGCGTCTGGCTGGCGCTCCAGAACGCGCCGAAGCCCGCGCTGGAGGTC  
CGCGGGCTCCGGGTGAGCCCCCTGCCCTCCGGCCCGAGCTCGTGCACTT  
CGAGGTGCCCCCTCTGCTCTGGCCGGCGGACGACGGATCGGTCTGGGGC  
20 ACTTCGAGTTCCGGCGCGATCGCGTTCGACGAGGGCGCGCGCAAGGAGATC  
GCGGCCGCAATTCACGCACCTCGTCGACGCGGTGATCGCCCGGCCGGACGC  
GCCGGTGTGACGCTCGTGAGGGGCGCCCGCGCCGAGGCCGCGCGAGCGC  
AGGCCGCGCTCGGCGAGGCGTTTCGCCAGGGCGGCGACGGCGCGCCTCGGC  
CAGCTGCGGCGTCGCTCGGCGGGCGACCGGACGCCCCGCGAGTAGCGGTC  
40 AGCCCTCGGCGGCGGCCAGGCGCACGCGGAACGGCGCAGGGTAGCCGTGG  
25 ACGCGCGCATGGGGTCGATCGCGCTGGGGACGCCGGCCCGCAGCAGCTG  
CTTGATGGCGAGCGAGATGTGCAGGATGGCCACGTACTTGCCGTGGCACG  
TATGGATCCCTGACTCCCAGAACAGGTAGTTTCTGCTCGGCAGGCGCCCG  
45 GGCTGAACTGGTCCGGGGCGTCGATGTGCTCGTGGTGTGTCATCGCCGA  
GGCGCTGCAGGCCATCACCAGCGCGCCGGCGGGGACCTTCTCCTCGTGCC  
30

5

10

15

20

25

30

35

40

45

50

55

GCGTGCCACGCCCCGACCGTGTAGTCGCGCACGCAGAGGCTCGTGACGCCG  
GTCGACGGGGGACGGAAGCGCAGCCCTCCAGCACATAGCCGGTGATGGC  
GGCGTCGTCTCGACGTTCAACACGTTGAGCGCGTCGCGCAGGACGCGCG  
GGCGCTTCATCAGCTCGACCAGGGCGTTGACGATCGCGCCGCCGCTGAGA  
5 TCCACGCAGCCCATGAGCAGCCCCAGGATCACGTCGCGGATCCCCCTCGTC  
GCTCTCGTAGGTCTCGGGGACCGACTGCATGACCAGGTAGCGGTCCAGCA  
CCGAGGGTTGCTCTGGCGGGGGCGACTTGGCCAGCTGCTTCTTCCGCGCG  
GCGACGATCGCGTCGATCATCGGCAGCGCCTCCTGACGAGCGGCCCTCGC  
CGCCGCCACGGCCGTCGGGTCGTTGGTCGGGTTGAGGAAGATCTCGTTGA  
10 ACAGCGCGTGGGTCCACGCCACCACCTTCTCGGTGCGGATCTCGCCGACG  
CCGAGGTACCGGGCCATCGCGCCGGCCGGCACCCCTGAGCGCGTAGTCACC  
GGTGAGATCGAACGGCTTGTGACGCCGACCTTGGCGAGCAGCCGTTTCG  
CCTCGTCCACGACGATCTGACGGTAGCGGGGAGATCGGCGCGCGGGAAC  
CGGAGGCGCAGGAGCGACTTCTCGTGCTCGTACTTGGGCGAGTCGTTTAT  
15 CGCCAGGATGTTCTGGCCCACGTTCTCGACCAGCTTGGGCGCGATGTTGT  
CGACCGAGAAGACGTCGTTGGCGTTGAGGACCTCGACGACGTCGTTGTAC  
CGGGTCACGAGCGTGATGGCCGGGATGGAGAAGATGGGCTTCTCGCGCCG  
CAGCTGGCTGAGGAACGGGAGCGGCTCCTCCCTCAGCCACTTGAACACCA  
TGCCGGCCTCGATCTGCTTCCGCTTGACCGGATCGTTCTCGTGCGCCAGC  
20 GCGCTGTGGAGGGCCTGCAGGTAATCGAACGGCGGCGCCTTGGCAGCGTC  
CGCTCGTCCCTCTTCTTCGATGTGAATGCTCATGGGGAGAATTCCTTTCT  
CGCATGCCGATCAGATCGCGACGCTCTGGGGGACCATCGACGGGAGCAGG  
TACAGGTACGGCTGCTCGCGGGCGCGGTTGCGCTCGGTGATCTCCCGCTC  
GATGTGCGCCAGGCGCTCGCGGTAGCGAGCAAACGCCTGCTTCGCCGCCG  
25 GATCTTGCAACAGGCACTTCATCGTCATCAGGCTCTCGCTGGTCAGCGGC  
GGGCCCATGCTGAGCGAGCGGCTGAGCACCATCTGGCGGATGCTCTGCGC  
GCGCCCCGGGAGGCGCTCCAGCGGCTTGAAGTGGCGCTTCTCGCTTCCGT  
TCAGGACATCGCCGTAGGAGCGGTAGGTGGCAAAGTGAAGCTTCGGGATC  
CAGGTGTAATAGTCCGTCTGCCCCAAGTTACGGCGGCGTGATATGCGGT  
30 CGCCGTGAAGATGATGTTGGTGACGATCGCGATCAGGTCGTGAGGCTCG

5

10

15

20

25

30

35

40

45

50

55

TGAGCTTCTCGAGCTGATCGGCTCGCTCCGGCGGGAGGAGGCTATCCATG  
CCGCCGAGCTGGGGGGACACGAGCTCGTGGATCCACCGCTGCAGGCTGGC  
GTCGCTCGACAGAGACCCCGGCGTCGGGTAGGCGATCTTCAGCACCTGTC  
CGACGTACTCCTGGATCGCGTCCCAGTGCAGCAGCGCGTCGTCGCGGTAG  
5 TGATAGCCGACCAGGTCGCGGACGTGCGCGCGCCGACAGGTCGCGGGGGAG  
CGCGCTCTCGTAGAACC GCCACGGCTTGCCGCCGTACCCTTTGATGCCCT  
15 TGCCGGTGTAGGCGCGCGTCAAGAGCTCGAACGAGCCCATGGTGGCCACC  
GAGCTCGTGATGTGGAAGAAGCGCCCTCGCCCGAGGAAGCGCCGGCGAGC  
CAGCTCGTTGATGGCCAGGGTGTTGAAGAAATGCGGCCTGAGCAGCTGGT  
20 10 GGAGCGGATGCGTCGCGGGCAGGTTGCGGTAGGTGCTCACCGCGAACGGC  
TCCACGATCAGGTGCGCGTACAGCAGGTGGGTACCTGGCCCTGGTAGAT  
GGCGTCGGCGCTCGCGACGGCGATCTTCGCCGTGAGCCAGTCGTCCGACG  
GACCCGAAGGGGTGAAGATCTTGTCGGGATGCGCCCCCTTCCCGGGGCGC  
25 GAGTGCACCAGCCTGATGGCCACGGGCAAGAGCTCACCGGCCGCGGTCTG  
15 GTGCAGCATGCACGTGCGCGCCAGCGGGTACTTGCCCAGCTCTTCCTGCA  
CGTCGGTGTGACGATGTCCTTGAAGATGCGGTAGTCGAGGAAGTAGAGC  
TGCCCGCCCTCGCGCACCTCCTCCAGCGTGCAGCCGTGCGCGATCGCGAT  
30 CGGCTTGGGCTCGGCGCCGCTCACGAAATCGGCGAGATCGGCCGGGTGCG  
CGCGCGGATGTGCGCCGGGTTGATCCCCACGAGGCGCTGCCGCCCGAAC  
20 TCGGCGTCCTCGGCCCAGCGCGTCCGACGAGGGGCTTGCGGATGAAGGT  
35 CCACGGCTTGAAGAACTCCTCGAACTGATCGAAGCTCTCCAGTTGTGCA  
TGGACTCGAAGATGGCGCCCAGCCGAGGTGCGACGTGGCCCTGAGGACG  
AACTTCCCTCGCGATAGCGCTTGTAACCGTACTCGAACAGGTGAAGCGC  
CTGCGCGATTTGCAGTCCGGCAGTGTCCTTCCACTTGCCGAGGTTGAGCG  
40 25 CCAAGATCTTCTTGATCGGAAAGCCTTCCCCCGGCGGTACGGAATCGCTG  
CCCGCCGGCAGGTTGGACCCGAAATTGCGCCAGTTCCGTGTGAGCTGGG  
CTCCCTCATGCTCGTGCTCGCTTCTCCGTCTCAGACGGACGGTGGATTGG  
45 GTGGTTACGTCAAACATCGCTCTCGCGTCGCAGCGGTCCGAGCGCGCGC  
CGGAATGGTTCCGTCTCAGTCGCAACAGGACTCAGTACATCCAGCGCCGC  
30 CCCCCGTCCTCGACCTGCCCCCGCAGCCGATCGCGCCGCCCTTCATCGTG



5

10

5

15

20

10

25

15

30

35

40

25

45

30

50

55

GAATCGACAGGTGCGATTCCACGAAAAGCCGCCGCGCCGAGTTGCACGCG  
ACCGATGCTCACGCGTGCAATTGTTGAGGCTGCTAGAAAACCGTGGAGCGT  
TCACGCATGTCAAGCCATTTTGTTCGGCGCCGCGGCGAGCGGCCGGATGC  
CGCGCGCCCCCGCGCCGGGCGTGTTCCGCTCCCGACGTACCGCTACCTCGA  
CGACGTATGGCTTGAAGGGCAACCGCGCAAGTCGTCCGATTCTGTGCTCGT  
ATCCTGCTCCTTCCAGCAGGATTTCCCCGCCGCCAGCGGCACAAAGGTGC  
CAGGGCGAGCAGAAAGAGCGCTGCCCGCCCCCTCCCCGGCCGCGCTCGCCT  
CGATCAGCGCGCGCTCGTTCGTTCGGTGCCTGAGACGTTTCGACGGACGTCAG  
TTAGTTAGCCTAGCTAACTTCAACACTGATGCGACTGATCGGGCCGACGC  
AACCGACGCAACCGACGCAACCGACGCAACCGACGCGACCGACGCAATCG  
ACGCAACCGACGTGACGGACGCTGGCGACTCGAAGAAAACCGGACGCA  
CTCCACGTCATCGACGTCATCGACGTCATCGATGCGCTCGATGCAATCCA  
TGCACTTGACGCGATCGGTGCGAGCAGGCGACGAGGTCCTCTCGTGAAAC  
ACCGAACCGAGTGCCGGTAGCGGGCGCGCCGCGAGTGTATGCTAGGCTCGG  
CCCTCTTGTCGAGGCCGCGCGCTCGGCGGTTCGAGCGTGGGCTCGGGTGCC  
GCGGTATCCGGCTGAACCAAGGAGGAGCGAGCCATGCAGGCAGATGACGA  
CGCGACGATCTACAAGGTGGTGGTGAACCACGAGGAGCAATACTCCATCT  
GGCCGGCGGACCGAGAGAACCCGCTCGGCTGGACGGAGGCCGCAAGACG  
GGCAACAAGGCGGAGTGTCTGGCGTACATCCAGGAGGTCTGGACGGACAT  
GCGCCCGCTCAGCCTCCGGAAGAAGATGGCCGAGAGCCCCCTGAATCGCGG  
CCCGCCCCGAGCGCCCGTCGCGAGCGGCCGGGCGGCGGGCTCAGCCGTGTC  
ATCGTCGCGCTCGACCGGCCGCGTCCCGCGGGATCGCGCGAGCCCGGCGG  
GGTCGTGCGCGCCGGCGCTTGTGCCGGGGCCCCCGCTCTCGTACGCCTCC  
GTCATGCCGCCCCCTCGATCTGCACGTCGCCTTGTTTCGGCGCCTCCGGCGC  
CGGCAAGACGGTCTCTCCTGGCAGCCTTCTACCGGGCGCAGACCCAGCCCT  
CGTTCCAGCAGGAGTACGCGTACAAGATCCAGGCGGTCAACAAGGCGCAG  
GGCAACCAGCTCCTCGGCCGGTTCTATCGCCTCGAAGAGGGCAGATTCCC  
GGACGGCAGCACGCGCTTCGACGAGTACGAGTTCGACTTCTTCCCGAGAG  
ATCTGCCCCGAGCCGGCGGTCCGCACTCCACTGGTACGACTACCCGGGACGC  
TGGTGGGAGGACGAGCCGGTCGACGCGGACGAGCGGGAGGCGATGCGCCA

5

10

15

20

25

30

35

40

45

50

55

GGGCCTCATCCGGCTCGGGATGAGCCAGGTGGGCATCCTCCTCGCGGACG  
GCGCGAAGTACCGGGCCGAGGGCACCGGGTACATCCGGTGGCTGTTGAG  
CACTTCGCCGACGAGTGCGACCGGCTGCGCCGGGCCAGCGCCGCCACGGG  
CGACGAGGTGAGCTTCCCGCGGGAGTGGATCCTCGCCCTCAGCAAGGCCG  
5 ATCTCTGCCCCCGGACTACAGCGCGCGGGACTTCGAGCGCGAGGTCTGC  
CGGGACGCCGACGATCAGCTGGCGAAGCTCTGCTCGGTGCTCCGCGCCGA  
GCACGCGTTTCGGCCACCGCTTCATGCTGCTCTCGTCGGTCGCGCCCCCGG  
CCGGCGCGCAGGTCGATCCGAGGACCTCGCTCGGCGTGCGCACCTCGCC  
CCCCGATCCTGGTGAGCACGGTCGAGGGCGCGGTGCGCGAGGCGCAGGC  
10 GGCGAGAAAGGAGAAGTCGGCCGGAAGACGTTCTTCCAGGGGCTGCGCG  
ATCTCGTGCAGTTTCGTCGACTCCCTCGACGACTTCCTGCCGAAGCGATAC  
CAGATCGTGAGCAAGATCCTGCGGTTTCATCTCGATCAAGGACTTCGCGAC  
CACCCGGCTCGACCGGCTCAAGAAGATGCGCGAGGACGCGATCCGGAAGG  
GCGACACCTTCACGGCGGTCTTGACCGCGATGGTCGCGGCCCTGCGCGAC  
15 GACGAGGGCGCCCCGCGCTACCACCAGAACCAGTGAGGTGTCATGCCCG  
CGCCAGCGCCCCTCGTCGAGACATCGCGCCTCCTCTGGAGGACGCGCGGC  
GAGCACTGGGATTACGAGTTTCATCTGTGTCCCCGAGATCCCGGCGCTGCC  
CGCCTGGCTCTCGACGCTCGAGGCGATGCTCGCCGACGCCGACGCCGGCG  
CCGGGGAGCTCCGCTATGGCCTGCTCGAGATCGACGATCGCGGGCAGAGG  
20 GCGCCGCGCGCCTATCCCTACGTGCGCGTGAGGTTCTTCGATCCGGCGCG  
GAGGGACTGGACCGGACGGCAGGTCCAGCACTTCGCGGCCTGGTTCCCGC  
CGGTCCCGCCCCGAGGCGGTGCGCGAGTTGCCAGAAGCGGTCCCGCCGAC  
TGGCACCTTCGCGTGCTCGACGGGCTCGCGGGGACGTACGGCTCCGGCGA  
GGTGTTTCGGGCTCCCCGAGGCGACGATCCGCGCCTGGAAGCGGAGCCACG  
40 ACGAGAGCCGGGCCGCGCGCGGATGGCGATCGTCAAGGCGACGCCGCCG  
GTTTCGCTGGGCGGCGGCGAGGCGCGCCGTGCGGTGGACGCGGGTGCC  
GACATTAAAAAAAAGCCGCCGAGCCGCCGGCCGCGGCGGGCCTCCTCT  
CGGTGGGCGCGGTCCCTAGCGGCCAGGGCCGGCGATTCCGCTGCTTCGCG  
ATCGGCGCCATGATGCTCGCCGCCCTTCTGTGACTGATGCTCGCTTGCGG  
30 TGTGCGCCTCCTCGGCGCCTGACGCGCTGCGCCGCGCAGGCCATCCGACGG

5

10

15

20

25

30

35

40

45

50

55

GGGGTCGGCCCGGCCAGCGCCCGCGGGCGACACCAGGGCATCGGCCCTC  
CGCTCGGGGCATCGATTGAGCTCTCCGAGCGGCGGTCCGTCGTCAATCGC  
CGCAGAGCTCCCACCGGGCGGAGCAGCTCTGGCCGGTGACCGCATAGGGG  
TTCGTCGGGCAGGTCCACCACTCGCCCTGGAAAGGACGCGGGTTGCAGTG  
5 CGGGAGGCACTCCACCCACCCCGACGAGCAGAGTTCCCTACCGAGACGG  
TCGGCTGAGCCGCGCAGAACCAGCGTTTTCTGCGAACCAGCCGGGATTG  
CACACATCGGGCGCCCCGCGACGGGCGGATACACCGTGGCGACCGCCTG  
GATGTCCACGGCGTCGAGGGTCTCGCTCCCGAAGCGCACGCGTCCTGAT  
CCGAACACCCACTCCAGTAGGTCATGATGGAGTCGTAGTCGTAGAAACCA  
10 GGGTTCACGACGATGTACCGCCGGCTCGAGGGCCACCCGCTGGCGACGTC  
GCTCGCGGGGAGCGGCTCTCGTTGGCTGCAGGCGCTCGGGACCAACGGGT  
GATGCCACTCATGCATGAAGCCGATCGCATGACCCATCTCGTGGATCGCG  
TACTGCTCCACGCAGTCGAAGCTGTATTCGACCCGGGCTGTCTGCCAGTT  
25 GTACTTGATGCAACGGTTGAAGTCGGCGCCCCAGGGCTTGAAGTGGACCG  
15 AGCCGCCCTTGTTGTAAACGCCGATCGAGTCCGATTGGTTGGGCGCGTCG  
GGGTGGATCCTGACGCCGACGTAGGTCATGCGAGTGGCCGGCAGGAGCGA  
ATCGCAGCTCTCCAGCCGGTGAAGCGAACCAGCTCCAGCGTTCCCAGC  
TGCCCTGGAGCGCGGTGCGCACGCGCGTGATGACGTCCGCGAGCGAGGGG  
TTGGGCGCATGGATCAGCCCGCCCGCGGCGCCGTGACCCCTCTGCTCCGC  
20 CGAGCTCGTGGGGTCGATGCAGACCGGGATCCGGACATGGCCGTGAGCGT  
35 CCTCAGGCCAGCGACTCGCGCTGTGGAAGACGCTCGCCTCGGCGGACCGC  
GGCGCGGCGGAGACGGTCAGCGCGGCGCCAGCGCTGCGAGGAGCAGCGG  
ACCGAGCGAAGAGCGAAACCGCACATGTCGTTTACGGGCCCCGCGTCGTGC  
GGTGACCCGAGACAATCTCGAGCGGGCTCATGGACGCAAACGCGTTGCGA  
40 TGGCCTTGACGATGTTCTTCTCCAATCGACGAGGGTTGTTCTGCTGAAC  
25 GCGGCTCCAGCGTGGAGCTCGACGCGGTTACCGGCTTCACGCCGGGGCC  
GTGGACGAGACCCGAGCACGGGGGAGGTTCGACCCGCACCGGCTCGCGGC  
GCCTCCACCCTGCACCTACGACGAGCCTGCCGCTCGGTTTCGCGGAAAAT  
45 GCCACCCCGCTGCCCAGCGGGCGAAGCGCGGACGAGGCGCTCGTCCCCAC  
30 GGTAGCGCCGGTGCCGCTGCATCCACCGCGCTCCTCCATGGGTGCGTGCC

55

5

10

15

20

25

30

35

40

45

50

55

CTCAGCGCCGGGGCCTACCTTATCGCATCTTGGGCGCTTGGCGTCCAGGA  
TTCGGCCTTAGACAGCACAAAGCAGAAGACCTTTGACACTGGATTTTTTCA  
TCATCGGCGGCGCTCGTTCTTCGCTGCGCCTCAAGCGCCGACCGTTCGTT  
TCGAAGCGAAGCGGTTTACGATCCGGGATCGCGGAAATTTGAAACGGAC  
5 GCGTCGCGCGGGCAACGCAGGGGACTCATCACGAGGCAACCGCGCTGCGT  
CGCGAAATTGGCCAGCCTCTCGGAGTCCCTAGTTCCGTGCGTCAGACGCG  
15 TCACCCACCATGTGAGCTCGGCGCGGCGCTCTACGTGCTTGAAAGACCT  
CCGCGAGCGCCGCGCTCTGCGCTCCGCGAGCGCACCAGGCTCCCCGTGGA  
TTCAGGGCAAGGCGGTCGTGATCACGTCCTTCGTGGTCGCCGTGCTCTGG  
20 CCGGGGCGGTGACGCGAGGACGGGTAGTACTGGCCGGCCCCGTTGAGCTT  
CTGGCAGACCCGGTCGCACGACCCGGCGATGCGGATCAGCCCGCACTCCA  
CGATCTGACCCCCGAGGTCACGTGGCCCGCGGCGCAGTCGCGCTGGTAG  
GCGCGCGAGTTGTGACCGTCGCGCTGTTGTAGCAAGCGTTGATATAGGG  
25 TTGCGCGGCGAACAGGTTGCCCCAGAAGGCCCCCTCGACGTCCGGATAAT  
15 CGATGAGCTCCTGGCTGGAGGAGAGCGTCTTCAGCGGATCCCGCAGCGAG  
CGGGCGGAGAGGAGCACCGGTACTTGATAGTAGTTCACGCGCGCCGCCAC  
GCAGCTGGACACGATGCGCTGCCCTGCGTCGTGAGCGGCCCCGCTCGCCC  
ACGCGGGCGCGACGCCGAGCAGCCGGGGTAGCGCTCGTCGTGCCTCTTG  
CCGTTTCGAGTCCGTCCACGAAAAATCGAAGGAGGCCGTGCTGCTCAGGGC  
20 GCAGCTCGCCGCGTAACGCAAGAAATCGCGCGCCAGCGCGCCGCTCGGCC  
35 CGGGATCCTGGATCGCGGCGAGGTTCCGCGCGCTGAGGCCGCTCAGGTTT  
AGGGCGTTTCAAGGTTTCAAGGCGTTGAGGTTTCAAGGCGTTGAGGTTTCAAGG  
GTTTCGTGCTGAGCGCGTTGCCGCCACGAGGGCCCCCTGGGATTCCCCCA  
CAGGCTCGCCCCACGCATCGGCGTCCACCACCTCGGCGGCGCAGCCCGAC  
40 AGCACCCCTGCCCAACCAAGCACGATGAATGTCCGCTCGAGAGACATGGA  
25 TTCCCCCGTGTTCTTGGCGCATGACCCGACGGCGCCCTGCGCGCGGCGCG  
CGCGGGCTCCCATCGATTGCTGGATGGGTTCAATATTCTACTTTTCCC  
GCGCTCTCGCGCCGGTGAAAGTCGCTTCAGCGGCGGCGAGGTCGATGTCA  
45 GGAGCGTCCGACTCCGTGCTCTCGTCAGCTCCGCGTACCAGCGACGGAG  
30 TCGCCCGCCCATGACGGTCGGAATGGTAGAGGCGGCCGCGAGGGCGCGCT

5

213

10

15

20

25

30

35

40

45

50

55

CGAGCTGCGCCCGGGCGTTCGGCGCGCCGGCCGCGGCGCAGGGCGGCGAGC  
GCCCCGCGCCTCGAGCACCTCGATCCGCTCCTGCCCCACCGAGCAGCGCGC  
GGAGCGCTCCTCGAGCGCGGCCACGCGGCGCGGTTCGTCGTCGCGGGTCG  
CGAGCTCGATCATCGCGCAGAGCACGTCTCTCCGAGGGCTTCAGCGCCTCG  
5 CAGCCGGCGTTCGTCTCGCGCCGCGCGGAGCCGCGAGCGCGATCCGGCGAGC  
GCCCCGCTCGTCGCCCTGGTAGAGCGCGAGGCGCGCGATCAGGAGCGTTA  
CGACGACCGGCGCGTGGCGATCGCCGCGAGCGCGGCGCTATCGCCTGGACG  
GCGCGCGCATGGGGCCTCGCGGCCCGAGATCGTCCATCAGGTACAGGTA  
CTCGGCGAGGTTGTAGCGGCCGACGAGCTCGAACGCGGGCTGGCCGAGCT  
10 CGCGCCCGAGCGCGATGGTTCGCTCGAAATCGGCGATCATCCCGGCGCGA  
TCGCCCTGGAGCGCCCGCGCGAGCCCGCGGTTGTTGAGCGCGGCGCCGAG  
GTGCATGAGATCGCTGCGCTCCTCGCAGCTGAGGATCACCGCGTCGAGGT  
CTCGCGCCGCTCCTCGACGCGGCGGAGGCTGGCCAGGATGAAGCCGAGC  
25 AGCAGCAGGGCGATGATGTGCGTCTCGTGGCCCTCGTCCCCGAGCCGCGC  
CGCCTGCGCCGCGGCGCGCGTTCAGCACCGCGGCGGCCTCGTCTCTCGCGGT  
CGGCGCGGTGGAGCGAGCGGCCACCGCGAGGAGCAGGCGGGCGCCGAGC  
AGGGCGAGGCCACCCGGCCGGCGAGGCGCTCGGCGGCCGCGACCCGCTC  
GCGCGCGGCCCGGTACTCGCCCGTCCAGTCGAGGATCATGGCCTCGTCGA  
GGAGGAGCTCGATCTCGGCCCGCCCTCCGACGCGCCGCGCCGCGCCTCG  
20 CGCGCCGCGGCGAGGTCGGCGAGGGCCTCGGTGTGGCGCCCGAGCCGAA  
GCGAGCGAGGCCCGCGCTCGGCGCTCCTCGGGGAGCAGCGCGCCGAGCA  
GCGCCTCGACGCGCCCGTAGCAGCCCTCGGCGTCGAGGTAGGCCCGGCGC  
GCGGCGCGAGCTCGGCGCCGCGGGCGAGGAGCGACCCGCGCGGGCGGT  
CAGGCCGCGCGCTCGCAGTTCGCGCCGCGAGCACCGCGGATCGGCCTCGC  
35 CCGCGGCCTCGAGCCAGTCGGCGGCGAGGCGGTGGCCGAGCGCGCGATCG  
GCAGCTCGCCGCGTAACGCAAGAAATCGCGCGCCAGCGCGCCGCTCGGCC  
CGGGATCCTGGATCGCGGCGAGGTTCCGCGCGCTGAGGCCGCTCAGGTTC  
AGGGCGTTTCAAGGTTCAAGGCGTTGAGGTTCAAGGCGTTGAGGTTCAAGGCG  
45 GTTCGTGCTGAGCGCGTTGCCGCCACGAGGGCCCCCTGGGATTCCCCCA  
30 CAGGCTCGCCCCACGCATCGGCGTCCACCACCTCGGCGGCGCAGCCCGAC

5

214

10

5

15

20

10

25

15

30

35

20

40

25

45

30

50

55

AGCACCCCTGCCCAACCAAGCACGATGAATGTCCGCTCGAGAGACATGGA  
TTCCCCCGTGTTCTTGGCGCATGACCCGACGGCGCCCTGCGCGCGGCGCG  
CGCGGGCTCCCATCGATTGCTGGATGGGTTCAATATTCTACTTTTTCCC  
GCGCTCTCGCGCCGGTGAAAGTCGCTTCAGCGGCGGCGAGGTCGATGTCA  
GGAGCGTCCGACTCCGTGCTCTCGTCAGCTCCGCGTACCAGCGACGGAG  
TCGCCCCGCCCATGACGGTCGGAATGGTAGAGGCGGCCGCGAGGGCGCGCT  
CGAGCTGCGCCCGGGCGTCGGCGCGCCGGCCGCGGCGCAGGGCGGCGAGC  
GCCCCGCGCCTCGAGCACCTCGATCCGCTCCTGCCCCGACCGAGCAGCGCGC  
GGAGCGCTCCTCGAGCGCGGCCACGCGGCGCGGTGCTCGTCGCGGGTCTG  
CGAGCTCGATCATCGCGCAGAGCACGTCTCCGAGGGCTTCAGCGCCTCG  
CAGCCGGCGTCTGTCGCGCCGCGCGGAGCCGAGCGCGATCCGGCGAGC  
GCCCCGCTCGTCGCTTGGTAGAGGCGCAGGCGCGCGATCAGGAGCGTTA  
CGACGACCGGCGCGTGCGCATCGCCGACGCGCGGCGCTATCGCCTGGACG  
GCGCGCGCATGGGGCCTCGCGGCCGCGAGATCGTCCATCAGGTACAGGTA  
CTCGGCGAGGTTGTAGCGGCCGACGAGCTCGAACGCGGGCTGGCCGAGCT  
CGCGCCCCGAGCGCGATGGTGCGCTCGAAATCGGCGATCATCCCGGCGCGA  
TCGCCCTGGAGCGCCCGCGCGAGCCCGCGGTTGTTGAGCGCGGCGCCGAG  
GTGCATGAGATCGCTGCGCTCCTCGCAGCTGAGGATCACCGCGTCGAGGT  
CTCGCGCCGCCTCCTCGACGCGGCCGAGGCTGGCCAGGATGAAGCCGAGC  
AGCAGCAGGGCGATGATGTGCGTCTCGTGGCCCTCGTCCCCGAGCCGCGC  
CGCCTGCGCCGCGGCGCGCGTACGACCGCGGCGGCCTCGTCCTCGCGGT  
CGGCGCGGTGGAGCGAGCGGCCACGCCGAGGAGCAGGCGGGCGCCGAGC  
AGGGGCGAGGCCACCCGGCCGGCGAGGCGCTCGGCGGCCGCGACCCGCTC  
GCGCGCGGCCCGGTACTCGCCCGTCCAGTCGAGGATCATGGCCTCGTCGA  
GGAGGAGCTCGATCTCGCCCCCGCCTCCGACGCGCGCCGCGCCGCTCG  
CGCGCCGCGGCGAGGTGCGCGAGGGCCTCGGTGTGGCGCCCGAGCCGAA  
GCGAGCGAGGCCCCGCGCTCGGCGCTCCTCGGGGAGCAGCGCGCCGAGCA  
GCCCTCGACGCGCCCGTAGCAGCCCTCGGCGTCGAGGTAGGCCCGGCGC  
GCGGCCGCGAGCTCGGCGCCGCGGGCGAGGAGCGACCGCGCGGGCGGT  
CAGGCCGCGCGCTCGCAGTGCGCCGCGAGCACCAGCGGATCGGCCTCGC

5

10

15

20

25

30

35

40

45

50

55

CCGCGGCCTCGAGCCAGTCGGCGGCGAGGCGGTGGCCGAGCGCGCGATCG  
TCCTTGGTGAGCTGCGCGTAAGCGCCCTCGCGCAGGAGCGCCTGGCGGAA  
GGAGTACTCCTCCTCGCCGGGGAAGCGGCCCTCGCGGTGGCGGACGCAGA  
GCTCCCCGGCGACGAGCGCGGAGAGGTGCTCCGCGAGCGGAGCGGCCTCG  
5 TCGCCCCCGAGCAGGTGCGCGACGGCGCCTCGCCAGAACACCTCGCCGAG  
CACGCTGGCGGCCCGCAGGATCCGGCGCGCGGGGGCGCGAGCGCCTCCA  
GCCGGACCTGCACCATCGCCACCACCGTCTCGGGCAGCGCGTCGCCGCGG  
CCCTCCGCCGTCGCGCGGATCAGCTCCTCGAGGAAGAACGGCTGGCCCTC  
GGACTGGGTGACCAGACGATCGATGAGGGCCCCGTCGGCCGCGTCGCCCA  
10 GCGCCTCCCGCGCGAGCTGCGCGCACGCCCTCGGCCGGAGCTGCCTGAGC  
CAGAGCTCCTGCCGCCCGCGCTCGGCCAGAGATCGGGGTACGCTTGC,

or their complementary strands,

15 (b) DNA-sequences which hybridise under stringent conditions  
to regions of DNA-sequences according to (a) encoding proteins  
or to fragments of said DNA-sequences,

20 (c) DNA-sequences which hybridise to the DNA-sequences accord-  
ing to (a) and (b) because of a degeneration of the genetic  
code,

25 (d) allele variations and mutants resulting by substitution,  
insertion or deletion of nucleotides or inversion of nucleotide  
segments of DNA-sequences according to (a) to (c), wherein the  
variations and mutants offer isofunctional expression products.

13. Peptide encoded by a DNA sequence according to claim 12  
selected from the group consisting of



5

Seq ID No 92

10

>Contig11\_002 591 amino acids MW=63639 D pI=5.80 numambig=0  
MLDVWSTSDQVACRLHCAGAGPSASLELRYDASAGARRDAERLAERLAALLEDLSRHPER  
5 PVAQGEYVGPGERAEIEAWSRGPAMELPSACALHRWFEERAEQHPDVVAVRSEGKSLTYG  
ELERRANRLASCLRRRGVGLDTIVGVCVPRSEDMVVATLAVLKVGGAYLPLDHEYPERL  
15 AFMMRDARARLLVTHDAIADELPTGGWTTLLLDAAEAEIAACSDARPAVSPPPDSGAYVI  
YTSGSTGTPKGSLSHRAIVNQMQWIIQRYWALTADDRVLLKAAFQFDVSVWEIFWPLSFG  
ARIVVARAGGHRDPEYLRRLVRDEGATTAYFVSSMLAAFLGGPEQPPASLRKVLVGGEA  
20 10 VPLDLVRRFYAKHDGDLINMYGPSEAAIAVTGCVLPSPRVTWVPLGAPVANAIEVFVLDG  
AMRRPAIGALGDLYIAGAPLARGYVGQPLTAERFLPDPCARAAGGRMYRTGDVARFLPD  
GMLEFQGRSDHQIKLRGHRIELGDVEAQIRRVPGVGQAQAVVLRDAPGDARLVAYVVLDG  
DAAGDAPDVRAGLKASLSAYMIPSSVRLYALPMCSERLAFTGSSYAGCLL\*

25

15 Seq ID No 93

30

>Contig11\_007 361 amino acids MW=38862 D pI=10.42 numambig=0  
MSDHEMTGFSLSPPQQRRAIRALDREAGAPGCRTLAVVAVTGPCDEGRLSAAALALAERHEI  
LRTRLVEAGRARPRRWSASRASRGRQDDWVGCSEAEQGERMSRLVARLSEDRGADDGLR  
VGLVRVGPEERRLVLAAPAWCVDEESIAPLVRELCASTAGAGAPPEQQYADVAEWLNGML  
20 ESEDAGDGRRFWAERRSHFGPPLHLAFSRGGAGAGAGSGRARVDLGQGGMAQVERWSSSW  
35 QVPQRIVLLALWASLLWRMSGGNEPEVTAVRFDGRSLDALAGAVGPFARFLPVRIEISA  
SDTLADVARRLALAEAEAAAHQDAAPGVSHRMSWGLLRGGRAGAVARRRAGPRARRLEH  
V\*

40

25 Seq ID No 94

45

>Contig11\_012 882 amino acids MW=95015 D pI=12.69 numambig=0  
MARALYAQEAAFRSAFNACAAAARAGLDLGEVVFREGECDGRTLLGTALAQPALFAVEL  
ALARLWMSWGIEPAAMIGHSLGELVAACVAGVFTLEDAMSLVIDRGRFMQAAPAGSMLAV  
GLPAADVEGLLEAGLEIAAVNSPKLTVVAGPASAIRDLAARLEAREVFARPLQTSHAFHC  
30 ALIDGAVAPFLESVRRARLSPPEIPVVSNTGALLTDAEATDPAYWARHLRQPVRFSQGV

50

55

5

10

15

20

10

25

30

35

40

45

50

55

EALFASGHALFLEVGPGRGLTTLVRQTLAGRGGAAIASLGSTHAASEPASLAEALGQLWE  
AGHAVDWTARPRGRPPARSSACRRTRSSGRGTGSRRAAAPPPPSRRRPTRPSRPRRTPSR  
RRRSPARTPGPRSPRRRPSPSARSPAIVERPPRRPRDRRPRLLRPARRLARRDPGDVPD  
PRAARRRAPAAGALPGADGRGARGPGRRRAPRRGAAARVPPDRADPPRRPAPAVVRAAPA  
5 VVRRSARAGQPGVQHPVRGARDGPARRRRAPPQVRDRAAPRGAADDVQRQGRRAVPGRR  
ARGPGAVPDVVRPRAPRGRGPRRGGLGARPRGEPRAVRPEPRAAPPRAGHPEAPRRARHRA  
RRAPRRVRRLVGRRLRRRARRALRRLRAGPAFFPAAGAPGAVRRLRRSPAGVAVRRGPGR  
APVLDDEALGRAAPGPGRGRPRAGGAKDVARGAAEPRRRGRAHPPDQGVLRAGGDLPLHG  
PARRVQARPAPAHRRARGPRRGHRRREPEPRRDRADDRLLRQPARAPDRLRGRPDVRRAGP  
20 PRARRRARGVRAPGPAVRPARRGPAAEGRRRARAPLRREVRHAQARPPDEARGARARGA  
GGRGDHDGVRLRPHGRRGRRLVPLRRRAQQRRAVPGGDGRQFP\*

Seq ID No 95

&gt;Contig11\_021 1213 amino acids MW=131017 D pI=12.40 numambig=0

15 MRGRRRRAAPHLRGARPARRGDRRGGLALRSRERARRRGHGEGLGAGGRRARRAARRR  
LPAPGPSPPRGAPQAPPGARRGPPRPHAVGGRDDRLARGHRAPRRRRGRAVARAARRAP  
AARRKHGRPGLRDLHVRVHGAAQGRDDRPPRRGEHGPRHQPPRRRPGGPGARALVAELR  
30 PVGLRRVRDARRRRRRRDPRPYRLGSGALARARGARAGDRVELGPGADGDAHGRVPRRG  
RPGAVVAPPRHDERRLDPAEAPRSHPRGLPRAPRREPRRRDRGVDLVDRPPDRGRRPGVA  
20 QHPLRPPAREPAHLRARRGAGAVSDRGPRRDPHRRDRRRARILARRGADPGAVPEAPHDR  
35 RAAVQDRRPGPLLRRRDHRAARAHPSGEDPRLPHRRARRDRGRPRAAPLGRAGGRGGEDR  
SVGREAPGRVRRRRRRRRWRAPRLRPEEAAGVHDPRGGRPPGAAERERQGGPRRPAGS  
RRGRAARRGRRPADRDRAAHLRGAGGLAGRGRRRRDRQPLRARLHLAAPRARAAPARRAH  
RGARARRGRRGAGRVADRPVPVPDHRATGAAARRGDGQGRAGRRRGAAARRGAPRRATPPG  
40 25 ARIAMSEPIETEDGGSDIAIVGMAGRFPGAPSVDALWENVRRGVESIARFPESEREPPV  
GASAAPGAPVVCAGGLDDIDRFDASYFGYSPREAQLMDPQQRLFLECAVALEDAGCDP  
ARFPGAIGVFGGCGSNTYLLQLLSHPDLAATVDPHALMLASEKDYLATRVSYKLDLHGPS  
45 VVVQTACSTSLVAVHMACESLLGGQCDLALAGGVSIGIPQKRGYPYVPGSICSPDGRCRP  
FDARAEGTVGGSGVGIVALKRLADALRDNTVHAVIRGSAVNNDGGRKVGFMAPSVDGQA  
30 AAISEAQSVAGVDPGSIGYVEAHGTATAIGDPPIEVEALTQAFRRKTPRKAYCALGSIKAN

5

10

IGHLDAAAGVAGLIKAAHVVRSGEIPPCVHFEAPNPKDLAASPPFFVPREAAPWPRELRP  
RRAGVSSFGIGGTNAHVVLEEPPPLPPRAPAPERDHLVLTLSARTPEALSTACAQLAAHLE  
ATDVPLDDVAFTLQTGRAEHPYRRVAVVARTRAEAIQGLAREGASALARPDEPRPSSRSRA  
RARRPSGWPARRSTRRRRRSGAPSTRARRRRGRAASISARSSSARARATGARCSAPRWRSR  
5 RSPSSSRSPGSG\*

15

Seq ID No 96

&gt;Contig11\_026 3079 amino acids MW=332984 D pI=5.97 numambig=0

20

10 MLTVVDHVVVEYWSFALIVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGG  
AEYWRKALDGATTAIDLPRDRARHDAGARRGRAHAITLPKPLTGALARLARERGTTLFSV  
LLSALTVLLHRASGQSDLVVGVPASGRHDEESARAFGYFVQMLPVRVALRGAASFDALVA  
RVRDAFLDALAHGDSALRHLLARAQGEAQRDALFDVAFQSTPPSLDARLSALAIGVGD  
VRIAQGELELTTLADEQAAAEFDLALFAAELDAGIALRFEYDQQLFDPATIERMARHFVV  
25 LLESAVEHPGRPLSELRLMSDAERALLDDWSGAAARQAASAPACVHALFEAHAARQ  
15 PDATALEFGHQRTYAQLSTWSTELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKA  
GAAYASLDPANPPARLAEMLADCRASLALTSSQASHKLTAAPCPVHLVQDGACAPSTHIP  
30 LVSRPDDLAYVLFTSGSTGTPKGVCVRHASLSRLVSFLHLRLDLSPSDRWTQVASSGFDA  
SVYEIWTPLACGAALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAEACFEQDWTGSA  
LRAMTVGGDKLHPLRRPLPFRLFNMYGPTEATVITVAEIADLGAEPPPLGRPVDLSALVYV  
20 LDPHMQPVPPGALGELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDLVRWRP  
35 DGQLAFAGRRDEQVKLRGRRVELGEVESVLRRLPGVREGIVVLHGQGSAAHLIAHVVPDA  
HPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKLLPAPPAHADYEPPSGE  
LELELAHIWQSVLHLDRVGRHDSFFDLGGHSLAMQVLGRIESSLGIRTTLRTLFEHPTL  
40 AQLASHLSSGAASTSAAAATALERGLTRPDGPSSPRVATPEEPFALTEGQRAMWLECQKS  
25 ADGALYNLGRTVRLGAGVDVAALRRAFEGLVERHEALRTTFLTRDGHPLQQVHRHVALEW  
AEEPAMALDEREIVARADEVRRRAFDLERGPLLRVHVWRRGEGQPPLLTVVVHHLVVVDYW  
SFALLVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGGAEYWRKALDGATTA  
45 IDLPRDRARHDASPRRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLHRASG  
QNCLVVGVPASGRNDDESTRAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDGLAHGD  
30 SALQHLLAEPRGAARRGGALFDVAFQALPSLDPRLAALTGAEDVRIAQGELELTTL

50

55

5

10

15

20

25

30

35

40

45

50

55

ADEQAAAEFDLALFAAELDSGIALRFEYDQQLFDPATIERMARHFVLLLESAVEHPGRPL  
SELRMLSDAERALLDDWSGAAARQAASAPACVHALFEAHAARQPDATALEFGHQRF  
TYAELSTWSTELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKAGAAAYASLDPANPP  
ARLAEMPLADCRAALVLTSSQASHKLTAAPCPVHLVQDGACAPSTH1PLVSRPDDLAYVLF  
5 TSGSSGTPKGVCVRHASLSRLVSFFQHLLALS PRDRWTQLASSGFDASVYEIWTPLACGA  
ALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAEACFEQDWTGIALRAMTVGGDKLHP  
LRRPLPFRLFNMYGPTEATVITTVAEVADLGDEPPLGRPIDSALVYVLDPHMQPVPPGVL  
GELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDLVRWRPDGQLAFAGRRDEQ  
VKLRGRRVELGEVESALRRLPAVREGVVVLHGQGSAAARLIAYVVPADPPSERDLREGMA  
10 RLVPDALVPAHFVLLPALPMSLSGKVDKLLPAPPAHADYEPSPSGELERELAHIWQSVL  
HLDRVGRHDSFFDLGCHSLLAMQVLGRIESSLGIRTTLRTLFEHPTLHQLADRLSSGAAS  
TTAAATVPASEIAPSLGRAPADEPYPLSYEQERLWVLEQLLPGGTAYNVVQAVRLRNLV  
DVDALSSALAALVRRHWSLRTVVFASPTPAQKICEPEAAPAEVVDLRGTPPDEAEAAARA  
25 WASREQATGFDLARGPVFRARLFRLDHDCVVLVLSHHI VTDASFQPLVRDLAELYRRA  
15 RGGGPADMPPELPLQYVDFAVWQRRHLAGKRLADKLAHWTATLRGLPVLELQTDPRPPVQ  
TFRGAERVLPDLARLVAQLDELARSRGATRFMVLLAALGVLLRRSSGQDDLAIGTAVANR  
PRPELEPLVGFFVNTIVMRLDLGGDPTFEELLSRARKVALEAFEHQDAPFEKVVEAVNPR  
30 RDLRSPLFQVMLVVQNAPTEALELGEVRIEPLDLPVEATRFDLRF SVEPRGGRDVISLQ  
YNVDLFDAATIDRMLATMQSVLSRATQDPAQVRVRLSVAPEDRERLVAWNDTAVATPDH  
20 LRLEEPFFERAVEQPDACAVVDAERRLTYGELARRAEAI AAAASRSGATANALVAVVMEK  
35 GWEQVA AVLGVLRAGAAYLPLDPRLP EERLRHLL EHA EVRLVLTQSAVDGTIAWPAGIER  
LAVDADERWREQPVARRPPGGSTDDLAYVIYTSGSTGLPKGVMIDHRGAVNTVLDINRRF  
DVGPEDRVLA LSSLSFDLSVYDVFGTLAAGGAVVIPDRTRASDPGHWR ELVERERVTVWN  
SVPALMEMLMDASPGAGDPALSSSLRLVMMSGDWIPLKLPDRIRAACRAPRVVSLGGATEA  
40 SIWSIAHPIADVDPAWRSIPYGRPLANQHTYVLDEGLEPCPIGVPGEIHIGGIGVALGYW  
25 RDEARTRERFLKHPTTGERLYRTGDLGRYFADGTIELLGRTDHQVKIRGFRIELGEIEAA  
LAQHPSVEQAVVAAKTDPSGEKRLVAYVVGADGDGAALRDFVRKKLPEYMI PAEVVVLPA  
45 LPLSANGKVDRAALPDPAAVAPRAAAVAPRTATERLIASVLAEVLQVEAVGVTDNLFELG  
FTSLLLVR AQRLLAERIAARAPDEGAAAQAVSLTDLFQYPTIEQLAQRLDAATVKAEPAD  
30 VGAQRAEARRDARRRRGRG\*

5

Seq ID No 97

10

&gt;Contig11\_011 544 amino acids MW=60164 D pI=9.10 numambig=0

MMSRIRAQLGVLEPLRALFQGPTVAALAAQVDAARRGEARRREFPPIARIPRDGPLPLSF

5 AQHRLWFVDQLEPGSPAYNIPFVVRATGRLDVDALRRSLFEIARRHEALRTTFSARDGVP

15

FPVVAPEARVPFRMSDLEHLAGEALDAVSAVLLEESLAPFDLSRGPLLVRVIRKRHDE

HVIALVVHHVFDVWSVGVFVGELAALYGGFAQQQPSRLPELPAQYVDFAAAQRAWLSGE

VLEGELRYWTTKLSGALRRARVPVDHEPAGRRTWRGARRSLDAGAELTRQIKAFCEAEI

SPFMALLAAYKLV LHQRTGLEDLVVGTDVANRNRVETEPMIGFFVNQLVLRDCCGGDPTF

20

10 GALVRRVRDVALEAFEHQDLPFDRLEALRPKCAVGHVPLFDAKFVMRNVHVPPMKLEGL

ELEALEGEATTTAFDFVLTVAEAGGSFRFGVEHSSELYRAATVDNFLSDYRQILATATAR

PDTPVSELRGELERAAAARRELERKAARGAALDKLTSARRRAVTLPRPGAPGEAKTSPKD

DLDE\*

25

15 Seq ID No 98

&gt;Contig12\_001 514 amino acids MW=56145 D pI=8.82 numambig=0

30

PPAVRRYVADRRPEQLPALAPEEREAAAARRLSALGAAPPQVRRRGLTRAPLSYGQSRIYF

LEQLSPGKPLFNVPGAVRLRGPVDVARLSAAFGEIVRRHDALRTSIANVDGELLOQIAQPH

AGFALDVVTSTPEEAAELDRRLRAEAWRPFAIGAPLLRATLFRLAEDEHVLLVTMHHVV

20 SDDWSLGVILRELLALYAGRSLPPPRQLQVSDFAAWQREMVESGALDGQRAYWRERLRGLS

35

RASISAGGGAEAPSHDPSGAIEEIALSPDKAAALEALARREGATLFMVLLALLDLVIHAR

SGALDIAVGTPIANRNRPELEDVVGLLTNTLVIRVDLARAGAFRDVLARARVQALDAFAN

QDIPFDVVTQDLKQERDHAQHPLFRVWLALQNAKPALEVRGLRVEPLPLRPELVHFEVA

LLLWPADDGSVVGHFEFRRDRVDEGARKEIAAAFTHLVDAVIARPDAPVSTLVEGARAEA

40

25 ARAQAALGEAFARAATARLGQLRRRSAGDRTPRE\*

Seq ID No 99

45

&gt;Contig12\_009 582 amino acids MW=65555 D pI=8.72 numambig=0

MREPSSTPNWRNFGSNLPAGSDSVPPGEGFPIKKILALNLGKWKDTAGLQIAQALHLFEY

30 GYKRYREGKFVLRATSDLGLGAIFESIDNWESFDQFEEFFKPWTFIRKPLVATRWAEDAE

50

55

5

10

15

20

25

30

35

40

45

50

55

FGRQRLVGINPAHIRRATPADLADFVSGAEPKPIAIAADGRTLEEVREGGQLYFLDYRIFK  
DIVDTDVQEELGKYPLAPTCMLHQTAAGELLPVAILVHSRPGKGAHPDKIFTPSGPSDD  
WLTAKIAVASADAIYQGQVTHLLYAHLIVEPFAVSTYRNLPATHPLHQLLRPHFFNTLAI  
NELARRRFLGRGRFFDITSSVATMGSEFELLTRAYTGKGIKGYGGKPWRFYESALPRDLA  
RDVRDLVGYHYRDDALLHWDALQEIYVGQVLKIAIPTPGSLSSDASLQRWIHELVSQOLGG  
MDSLPPPERADQLEKLTSLDDLIAIVTNIIFTATAYHAAVNFQTDYYTWIPNAQFATYR  
SYGDVLNGSEKRQFKPLERLPGRAQSIQMVLSRSLMGPPLTSESLMTMKCLLQDPAK  
QAFARYRERLAHIEREITERNRAREQPYLYLLPSMVPOSVAI\*

10 SEQ ID No 100 (>ORF1)  
VSQRTSCYLRGGGVCSMNDAFLALERNERNRPSTVIDLLRQRAEAEAPARPIYCFLESGDVEAG  
ATWVTLREIDERARTVAALLQASGVAPGARALLYPGIEYITAFFGCLYAGVRTVPAYPPDL  
GRLERTLPRVASIVADARAEAAALTSSAVAGIVASLPASAAAAALQRLRWIATDGPSPGPIEGP  
GAALRPESVAFLOQYTSGSTGEPKGVMLTHGNLLHNSRLIAHGFDLTSPDPVGVWLPPYHDMG  
LIGGILQALYRRIRVALMSPLSFLQRPWRWLRAVSALGASVSGGPNFAYDLCVRKSSEEERAA  
LDLRSWEVAFTGAEPVRADTLDRFARAFVSGFRREAFYPCYGLAEATLIVSGGARAEAPVLA  
RLAPEEVELGRAVASAAEGARVFGVSGRALDPRVAIVDPAGNELGPGEIGEIVVSGPSVAVG  
YWGRPEETEATFGATLAGSAAAPRYLRTGDLGFLRGGEFVVGRSKDLIIILGRNHFQDIEKT  
VESSHRAVRPGCSAAFSVEHEGEERLAVVCEVDPRVAADPREIVAAREAVTAEHQLVAHAVAL  
IAPGALPKTSSGKVRRECRRAFLDALGERHVAFAPELLDDASPPDDAPPETEESGRSLLD  
ALRSTLARALRLDAGQIDDALPISRFGLDLAARELQHAQVQRTGRAIPLTSILRGGSLRLTR  
EITRLDGPSSPRVATPGGAVCADRWGTGRFGSSAISRPMERFTTWAGRSGSVPAFKRVDLRR  
F

25 SEQ ID No 101 (>ORF2)  
VYSSAYVLFAVCAGTTRVASAPETAGFPLECVGDDGTVLGPDSFVVGYTQVYVFKKERLNTNP  
PIDGFTLKLKGNEVAPGEDGLPVVKRCVRSEEQAQGCGRTEPAEDECTTYEIEAVVPEKAAEV  
DEEAAGLGGPPAREAIWVDYYTDGGEFDGARRLVSDTTGASRGNGTTWTPPSEPGRVSLWAV  
VHDTRGGASVTRREVQVE

5

SEQ ID No 102 (&gt;ORF3)

10

VVGTVLSAGTGEPLPDIAVTLVRPDGGREEAKTDQGGKFRFKNLPPGKYRVEVAAAGFEPFAA  
EEEIAAGEAIEVRYRISLAAPQDGKAPGIEVTVQGERPPREVTRRTIERREIDRIPGTGGDAL  
RSLQSLPGVARSGFGLLIVRGSAPQDTLTFVDRTVPVPIIYHFGGLSSVVPTMLEKIDFYPGN

5

FSAVYGRAMGGIVDVGLRSPKQDGKYHGVVQLDLIDGRVLLEGVPVFLKDWTFFIAAGRRSWVD

15

AWLGPVLKEAGSSVTQAPVYYDYQFVLEGRPSASERVSRASYGSDDAFKITLTKPPEDEPAL

GDFGLHTAFQRFQLSYENRIGSRDRLLWSMALGRDIADFEISPLAFNVVSTSLDLRLELSHRF

ARYLTMNVGTDLSGGVATVNIRAPSQQPAGHPSNQPFSTYPFQDRSFDGAYSRAAYAELEV

20

10 GNAELKSNRAVHYGLGVEQEITPQIEVTLDFYKQLDRLVVFSSPEKDDYADGTGYAVGGELLL

KYPDERFFGWAAAYTLRSVRKDGPDDEEHLTQFDQTHVLTVLGSLRLGRGWELARFRLVSGN

LQTPYVCDPEEKGCNPNRVNAIYHASSARYSPIPLGGDYSERMPLFHQLDIRADKTWKFKRWQ

25

LGLYLDIQNVYNYMAAEGISYNFNYTKREYVTGLPFLPTLGLRGDF

15 SEQ ID No 103 (&gt;ORF4)

VIAVDNNPEAVDAVKDKTSAAFGDATVHKVLEGIGAQYVETAIVTFGEHFEPVLCVASLVR

30

MGVRIIARAATDRQADILRAVGATRVIQLETEMGRRVGADITMPLAQDLLDLASHYRVVPWNA

HGPLVGQTLAGSKIRQRYRINVLGVRPHTNKRPGDKPRLEAPTPDYVIRDGDTLLLVGDSDDV

SRFVAEVGG

20

35

SEQ ID No 104 (&gt;ORF5)

SGSSGGGSSAEGSRCQPSGGGPHWLLEGETVTFPVTCASGLALAGDAFEVGPLPEGAAYDPIA

REVTFSPLDQAAVYDIEIRVAQTSEVGRVKVGADAFADPSNVPVVDPTYPEEYGLPVLFL

40

SPVPEDKEYAPATVIYRGHTYAAEAELRGESSLSYPKRSYTLKFPKDDKFNEPDEAGGFTDRR

25

KVVLTTFDDNSYVRQRLAYDLWNRLDPEHIQIKTYSAVLYLDGEYAGLYTVADHVDGYLMED

HGYPQDGNLYKAVSHDANFALTDRSGDPKDTLHDGFEEKKEGAPAEGEPEAFSDLEDLVSFVAE

SDDATFAAEIGSRIDLRYEDWWIFVTFIVANDSAGKNSYHYRDPADGVFRYAPWDFNASFG

45

QSWETEREPASDRVDYRDVNLLFERLLEEPSIGDPLRARYDQVLRGALAEAEIHAIVDGYVER

IDASARRDEARWGEAYRSYEGWSWRDDFTTYEEIAYLK

30 AWISERWQHQDELY

50

55

5

10

SEQ ID No 105 (Contig 11 >ORF1)

VLDVWSTSDQVACRLHCAGAGPSASLELRYDASAGARRDAERLAERLAALLEDLSRHPERPVA  
QGEVGPGERAEIEAWSRGPA MELPSACALHRWFEERAEQHPDVAVRSEGKSLTYGELERRAN  
5 RLASCLRRRGVGLDTIVGVCVPRSEDMVVATLAVLKVG GAYLPLDHEY PGERLAFMMRDARAR  
15 LLVTHDAIADELPTGGWTTLLLDAAEAEIAACSDARPAVSPPPDSGAYVIYTS GSTGT PKGSL  
ISHRAIVNQMWIQRYWALTADDRVLLKAAFGFDVSVWEIFWPLSFGARIVVARAGGHRDPEY  
LRLVRDEGATTAYFVSSMLAAFLGGPEQFPASLRKVLVGGEAVPLDLVRRFYAKHDGDLIN  
MYGPSEAAIAVTGCVLPSPDRVTWVPLGAPVANA EVFVLDGAMRRPAIGALGDLYIAGAPLAR  
20 10 GYVGQPGLTAE RFLPDPCARAAGGRMYRTGDVARFLPDGMLEFQGRSDHQIKLRGHRIELGDV  
EAQIRRVPGVGQA AVVLREDAPGDARLVAYVVL DGDAGDAPDV RAGLKASLSAYMI PSSVVR  
LYALPMC SERLAFTGSSYAGCLL

25

SEQ ID No 106 (Contig 11 >ORF2)

15 MSDHEMTGFSLS PQQRAIRALDREAGAPGCRTLAVVA VTGPCDEGRLSAAALALAERHEILRT  
RLVEGRARPRRWSASRASRG RQDDWVCSEAEQGERMSRLVARLSEDRGADDGLRVGLVRVG  
30 PEERRLVLAAPAWCVDEESIAPLVRELCASTAGAGAPPEQQYADVAEWLNGMLESEDAGDGRR  
FWAERRSHFGPPLHLAFSRGGAGAGAGSGRARVDLQGGMAQVERWSSSWQVPQRIVLLALWAS  
LLWRMSGGNEPEVTVA VRFDGRSLDALAGAVGP FARFLPVRIEISASDTLADVARRLALAEAE  
20 AAAHQDAAPGVSHRMSWGLLRGGGRAGAVARRRAGPRARRLEHV

35

40

SEQ ID No 107 (Contig 11 >ORF3)

MSRIRAQLGVELPLRALFQGPTVAALAAQVDAARRGEARRREFPPIARI PRDGPLPLSFAQHR  
LWFVDQLEPGSPAYNIPFVVRATGRLDVDALRRSLFEIARRHEALRTTFSARDGVFPFVVAPE  
25 ARVPFRMSDLEHLAGEALDA AVSALVLEESLAPFDLSRGPLL RVRVIRKRHDEHVIALVVHHV  
VFDVWSVG VFGELAALYGGFAQGQPSRLPELPAQYVDFAAAQRAWLSGEVLEGE LRYWTTKL  
SGALRRARVPVDHEPAGRRTWRGARRSLDAGAE LTRQIKAF CEREAI SPFMALLAAYKLVLHQ  
45 RTGLEDLVVGTDVANRNRVET EPMIGFFVNQLVLR TDCCGDPTFGALVRRVRDVALEAFEHQD  
LPFDRLVEALRPKGAVGHVPLF DAKFVMRNVHVPPMKLEGLELEALEGEATTTAFDFVLTVAE

50

55



5

AGGSFRFGVEHSSELYRAATVDNFLSDYRQILATATARPDPVSELRGELERAAAARRELERK  
AARGAALDKLTSARRRAVTLPRPGAPGEAKTSPKDDLDE

10

SEQ ID No 108 (Contig 11 >ORF5)

5 MSEPIETEDGGSDIAIVGMAGRFPAGPSVDALWENVRRGVESIAARFPESEREPPVGASAAPG  
15 APVVCAGGLLDDIDRFDASYFGYSPREAQLMDPQQRLFLECAVALEDAGCDPARFPGAIGVF  
GGCGSNTYLLQLLSHPDLAATVDPHALMLASEKDYLATRVSYKLDLHGPSVVVQTACSTSLVA  
VHMACESLLGGQCDLALAGGVSIGIPQKRGYPYVPGSICSPDGRCPFDARAEGTVGGSGVGI  
VALKRLADALRDRNTVHAVIRGSANNDGGRKVGFMAPSVDGQAAAISEAQSVAGVDPGSIGY  
20 10 VEAHGTATAIGDPIEVEALTQAFRRKTPRKAYCALGSIKANIGHLDAAAGVAGLIKAAHVVR  
GEIPPCVHFEAPNPKLDAASPFFVPREAAPWPRELRPRRAGVSSFGIGGTNAHVVLEPPPL  
PPRAPAPERDHLVTLTSAARTPEALSTACAQLAAHLEATDVPLDDVAFTLQTGRAEHPYRRVVA  
RTRAEAIQGLAREGASALARPDEPRPSSRSRARARRPSGWPARRRRSGAPSTRARRRRGR  
25 AASISARSSSARARATGARCSAPRWRSPRSSPSSSRSPGSG

15

SEQ ID No 109 (Contig 11 >ORF6)

30 VVDHHVVVEYWSFALIVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGGAEYWRK  
ALDGTTAIDLPRDRARHDAGARRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLH  
RASQSDLVVGVPSAGRHDDESARAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDALAHG  
20 DSALRHLLARAQGEAQRDALFDVAFQSTPPSLDARSALAIIGVDVRIAQGELELTTLADEQ  
35 AAAEFDLALFAAELDAGIALRFEYDQQLFDPATIERMARHFVVLLSAVEHPGRPLSELRLMS  
DAERALLDDWSGAAAARQAASAPAPACVHALFEAHAARQPDATALEFGHQRTYAQLSTWST  
ELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKAGAAYASLDPANPPARLAEMLADCRAS  
40 LALTSSQASHKLTAAPCPVHLVQDGACAPSTHPLVSRPDDLAYVLFTSGSTGTPKGVVVRHA  
25 SLSRLVSFLHLRLDLSPSDRWTQVASSGFDASVYEIWTPLACGAALLLADDDALRSPTALVSW  
LVAQRATLSFMPTPLAEACFEQDWTGSALRAMTVGGDKLHPLRRPPFRLFNMYGPTEATVITT  
VAEIALDLGAEPPLGRPVDSALVYVLDPHMQVPPGALGELYIGGACLAQGYTRTDLTAEERFLP  
45 DPFQPGARLYRTGDLVRWRPDGQLAFAGRRDEQVKLRGRRVELGEVESVLRRLPGVREGIVV  
LHGQGSAAHLIAHVVPDAHPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKLLP  
30 APPAAHAJYEPPSGELELELAHIWQSVLHLDRVGRHDSFFDLGGHSLAMQVLGRIESSLGIR

50

55

5

TTLRTLFEHPTLAQLASHLSSGAASTSAAAATALERGLTRPDGPSSPRVATPEEPFALTEGQR  
AMWLECQKSADG

10

ALYNLGRTVRLGAGVDVAALRRAFEGLVERHEALRTTFLTRDGHPLQQVHRHVALEWAEEPAM  
ALDEREIVARADEVRRRAFDLERGPLLRVHVWRRGEGQPPLTVVVHHLVVDYWSFALLVREL

5

GELYSALRAGRPPQLPPSSFFAAGVSCSPREAAAGGAEYWRKALDGATTAIDLPRDRARHDA

15

SPRRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLHRASGQNDLVVGVPSAGRND  
DESTRAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDGLAHGDSALQHLLAEPRGAARRGG  
ALFDVAFAFQGALPSLDPRLAALTGAEDVRIAQGELELTTLADEQAAAEFDLALFAAELDSG  
IALRFEYDQQLFDPATIERMARHFVLLLESABVEHPGRPLSELRLMSDAERALLDDWSGAAAA

20

10

RQAASAPAPACVHALFEAHAARQPDATLEFGHQRTYAELSTWSTELALWLRDRGVGPGSVV  
GVCIERSPRMVAAQLAVLKAGAAAYASLDPANPPARLAEMLADCRAALVLTSSQASHKLTAAPC  
PVHLVQDGACAPSTHIPLVSRPDDLAYVLFTSGSSGTPKGVCRHASLSRLVSFFQHLLALS  
RDRWTQLASSGFDASVYEIWTPLACGAALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAE  
ACFEQDWTGIALRAMTVGGDKLHPLRRPLPFRLFNMYGPTTEATVITTVAEVADLGDEPPLGRP

25

15

IDSALVYVLDPHMQPVPPGVLGELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDL  
VRWRPDGQLAFAGRRDEQVKLRGRRELGEVESALRRLPAVREGVVVLHGQGSAAARLIAYVVP  
GADPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKLLPAPPAHADYEPPSGEL  
ERELAHIWQSVLHLDVRGRHDSFFDLGGHSLAMQVLGRIESSLGIRTTTLRTLFEHPTLHQLA  
DRLSSGAASTTAAATVPASEIAPSLGRAPAD

30

20

EPYPLSYEQERLWVLEQLLPGGTAYNVVQAVRLRNLDVDALSSALAALVRRHWSLRTVFNAS  
PTPQKICEPEAAPAEVVDLRGTPPDEAEAAARAWASREQATGFDLARGPVFRARLFRLDHVC  
VLVLSTHHIVTDAWSFQPLVRDLAELYRRARGGGPADMPPELPLQYVDFAVWQRRHLAGKRLAD  
KLAHWTATLRGLPVLELQTDPRPPVQTFRGAERVLPDLARLVAQLDELARSGATRFMVLLA  
ALGVLLRRSSGQDDLAIGTAVANRPRPELEPLVGFFVNTIVMRDLGGDPTFEELLSRARKVA

35

25

LEAFEHQDAPFEKVVEAVNPRRDLRSPLFQVMLVVQNAPTEALELGEVRIEPLDLPVEATRF  
DLRFSVEPRGGRDVISLQYNVDLFDAAATIDRMLATMQSVLSRATQDPAQRVRALSVAPEDRER  
ALVAWNTAVATPDHLRLEPPFERAVEQPDACAVVDAERRLTYGELARRAEIAAAAASRSGA  
TANALVAVVMEKGWEQVAAVLGVLRAGAAYLPLDPRLPEERLRHLLHAEVRLVLTQSAVDGT  
IAWPAGIERLAVDADERWREQPVARRPPGGSTDDLAYVIYTSGSTGLPKGVMIDHRGAVNTVL

45

30

DINRRFDVGPEDRVLALSSLSFDLSVYDVFGTLAGGAVVVPDRTRASDPGHWRLEVERERV

50

55

5

10

VWNSVPALMEMLMDASPGAGDPALSSRLVMMSGDWIPLKLPDRIRAAACRAPRVVSLGGATEA  
SIWSIAHPIADVDPAWRSIPYGRPLANQHTYVLDEGLEPCPIGVPGEIHHIGGIGVALGYWRDE  
ARTRERFLKHPTTGERLYRTGDLGRYFADGTIELLGRTDHQVKIRGFRIELGEIEAALAQHPS  
VEQAVVAAKTDPGSEKRLVAYVVGADGDGAALRDFVRKKLPEYMI PAEVVVL PALPLSANGKV  
5 DRAALPDPAAVAPRAAAVAPRTATERLIASVLAEVLQVEAVGVTDNLFELGFTSLLLVRQRL  
15 LAERIAARAPDEGAAAQAVSLTDLFQYPTIEQLAQRLDAATVKAEPADVGAQRAEARRDARRR  
RGRG

20

10 SEQ ID No 110 (Contig 12 >ORF1)  
PPAVRRYVADRRPEQLPALAPEEREAAAARLSALGAAPPQVRRRGLTRAPLSYGQSRIYFLEQ  
LSPGKPLFNVPGAVRLRGPVDVARLSAAFGEIVRRHDALRTSIANVDGELLQIAQPHAGFALD  
VVTSTPEEAAELDRRLRAEAWRPFAIGAPPLLRLATLRLAEDEHVLLVTMHHVVSDDWSLGI  
LRELLALYAGRSLPPRLQVSDFAAWQREMVESGALDQRAYWRERLRGLSRASISAGGGAEA  
25 PSHDPSGAIEEIALSPDKAAALEALARREGATLFMVLLALLDLVIHARSGALDIAVGTPIANR  
15 NRPELEDVVGLLTNTLVIRVDLARAGAFRDVLARARVQALDAFANQDIPFDVVTQDLKQERDH  
AQHPLFRVWLALQNAPKPALEVRGLRVEPLPLRPELVHFEVALLWPADDGSVVGHFERRDR  
30 VDEGARKEIAAAFTHLVDAVIARPDAPVSTLVEGARAEAAQAALGEAFARAATARLGQLRR  
RSAGDRTPRE

35

40

20 SEQ ID No 111 (Contig 12 >ORF2)  
MSIHIEEEGRADAAKAPPFDYLQALHSALAHENDPVKRKQIEAGMVFKWLREEPLPFLSQLRR  
EKPIFSIPAITLVTRYNDVVEVLNANDVFSVDNIAPKLVENVGQNILAMNDSPKYEHEKSLLR  
LAFPRADLPRYRQIVVDEANRLAKVGVDKPFDLTGDYALRVPAGAMARYLGVGEIPTKVVA  
WTHALFNEIFLNPTNDPTAVAAARAARQEALPMIDAIVAARKKQLAKSPPEQPSVLDRYLVM  
40 25 QSVPETYESDEGIRDVILGLLMGCVDLSSGAIVNALVELMKRPRVLRDALNVVNVEDDAAITG  
YVLEALRFRPPSTGVTSLCVRDYTEVGRGTRHEEKVPAGALVMACSSAMHDHEHIDAPDQFRP  
GRLPSRNYLFWESGIHTCHGKYVAILHISLAIKQLLRAGVPSAIDPMPRVHGYPPAPFRVRLAA  
45 AEG

50

55

5

SEQ ID No 112 (Contig 12 &gt;ORF3)

10

MREPSSTPNWRNFGSNLPAGSDSVPPGEGFPIKKILALNLGKWKDTAGLQIAQALHLFEYGYK  
RYREGKFVLRATSDLGLGAIFESIDNWESFDQFEEFFKPWTFIRKPLVATRWAEDAEFGRQRL  
VGINPAHIRRATPADLADFVSGAEPKPIAIADGRTLEEVREGGQLYFLDYRIFKDIVDTDVQE

5 ELGKYPLAPTCMLHQTAAGELLPVAIRLVHSRPGKGAHPDKIFTPSGPSDDWLTAKIAVASAD

15

AIYQGQVTHLLYAHLLIVEPFAVSTYRNLPATHPLHQLLRPHFFNTLAINELARRRFLGRGRFF  
DITSSVATMGSEFELLTRAYTGKGIKGYGGKPWRFYESALPRDLSARDVRDLVGYHYRDDALLH  
WDAIQEYVGQVLKIAYPTPGSLSSDASLQRWIHELVSQOLGGMDSLPPERADQLEKLTSLDD  
LIAIVTNIIFTATAYHAAVNFGQTDYYTWIPNAQFATYRSYGDVLNGSEKRQFKPLERLPGRA

20

10 QSIRQMVLRSRLSMGPPLTSESLMTMKCLLQDPAKQAFARYRERLAHIEREITERNRAREQP  
YLYLLPSMVPQSVAI

SEQ ID No 113 (Contig 12 &gt;ORF4)

25

VSSSRSTGRVPRDRASPAGSCAPALVPGPPLSYASVMPPLDLHVALFGASGAGKTVLLAAFYR

15 AQTQPSFQQEYAYKIQAVNKAQGNQLLGRFYRLEEGRFPDGGSTRFDEYEFDFPRDLPEPAVR

IHWYDYPGRWWEDEPVDADEREAMRQGLIRLGMSQVGILLADGAKYRAEGTGYIRWLFEFHAD

30

ECDRLRRASAATGDEVSPREWILALSKADLCPPDYSARDFEREVCRDADDQLAKLCSVLRAE

HAFGHRFMLLSSVAAPAGAQVDPRTSLGVRTLAPAILVSTVEGAVREAQAARKEKSAGETFFQ

GLRDLVQFVDSLDDFLPKRYQIVSKILRFISIKDFATTRLDRLKMRDAIRKGDFTTAVLTA

20 MVAALRDDEGARAYHQNQ

35

SEQ ID No 114 (Contig 12 &gt;ORF5)

MPAPAPLVETSRLLRTRGEHWYEFICVPEIPALPAWLSTLEAMLADADAGAGELRYGLLEI

40

DDRGQRAPRAYPYVAVRFLDPARRDWTGRQVQHFAAWFPPVPEAVAELPEAVPADWHLRVLD

25 GLAGTYGSGEVFGLEPEATIRAWKRSHDESRAARAMAIVKATPPVSLGGGEAAPSRWTRVPTLK

KKPPEPPAAAGLLSVGAVPSGQGRRFGCFAIGAMMLAFCRLMLACGVRLIGA

SEQ ID No 115 (Contig 12 &gt;ORF6)

45

VRFRSSLGPLLLAALGAALTVSAAPRSAEASVFDSASRWPEADGHVRIPVCIDPTSSAEQRV

30 DGAAGGLIHAPNPSLADVITRVRTALQGSWERWSSVRFTGWESCDLLPATRMTYVGVRIHPD

50

55

5

10

APNQSDSIGVYNKGGSVQFKPWGADFNRCIKYNWQTARVEYSFDCVEQYAIHEMGHAIGFMHE  
WEHPLVPSACSQREPLPASDVASGWPSSRRYIVNPGFYDYDSIMTYWSGCSDQDGVRFGSET  
LDAVDIQAVATVYPPVGGAPDVCNPGWFAGKRWFCQAQPTVSVGNSCSSGWVECLPHCNPRPF  
QGEWWTCTNPYAVTGQSCSARWELCGD

5

15

SEQ ID No 116 (Contig 12 &gt;ORF7)

VGESQGALVGGNALSTNALNLNALNLNALNLNLSGLSARNLAAIQDPGPGALARDFLRY  
AASCALSSTASFDTSNGKRHDERYPGLLGVAWASGGLDDAGQRIVSSCVAARVNYQ  
VPVLLSARSLRDPLKTLSSSQELIDYPDVEGAFWGNLFAAQPYINACYN SATVDNSRAYQRDC

20

10 AAGHVTSGGQIVECGLIRIAGSCDRVCQKLNAGQYYPSCVDRPGQSTATTKDVITTALP

SEQ ID No 117 (Contig 12 &gt;ORF8)

25

VLAHCERGGLTARAASLLARGAELAAARRAYLDAEGCYGRVEALLGALLPEERRARGLARFR  
LGRHTEALADLAAAREAAAAASEAGAEIELLLDEAMILDWTGEYRAARERVAAAERLAGRVAS

15

PLLGARLLLGVGRSLHRADREDEAAAVLTRAAAQAARLGDEGHETHI IALLLLGFILASLGRV  
EEAARDLDAVILSCEERSDLMHLGAALNNRGLARALQGDRAGMIADFERTIALGRELGQPAPFE

30

LVGRYNLAEYLYLMDDLAAARPHARAVQAIAPRCGDRHAPVVVTLLIARLRLYQGDEAGARRI  
ALRLRAARDDAGCEALKPSEDVLCAMIELATRDDRAAWAALEERSARCSVGQERIEVLEARA  
LAALRRGRRADARAQLERALAAASTIPTVMGGRLLRWYAELTRATESDAPDIDLAAAEATFTG

20 ARAREKVEY

35

SEQ ID No 118 (Contig 12 &gt;ORF9)

QAYPDLWAERGRQELWLRQLPPRACAQLAREALGDAADGALIDRLVTQSEGQPPFLEELIRAT  
AEGRGDALPETVVAMVQVRLEALAPPARRILRAASVLGEVFWRGAVAHLLGGDEAAPLAEHLS

40

25 ALVAGELCVRHREGFRPGEEYSFRQALLREGAYAQLTKDDRALGHRLAADWLEAAGEADPLV  
LAAHCERGGLTARAASLLARGAELAAARRAYLDAEGCYGRVEALLGALLPEERRARGLARFRL  
GRHTEALADLAAAREAAAAASEAGAEIELLLDEAMILDWTGEYRAARERVAAAERLAGRVASP

45

LLGARLLLGVGRSLHRADREDEAAAVLTRAAAQAARLGDEGHETHI IALLLLGFILASLGRVE  
EAARDLDAVILSCEERSDLMHLGAALNNRGLARALQGDRAGMIADFERTIALGRELGQPAPFEL

30 VGRYNLAEYLYLMDDLAAARPHARAVQAIAPRCGDRHAPVVVTLLIARLRLYQGDEAGARRIA

50

55

5

10

LRLRAARDDAGCEALKPSEDVLCAMIELATRDDDRAAWAALEERSARCSVGQERIEVLEARAL  
AALRRGRRADARAQLERALAAASTIPTVMGGRLRRWYAELTRATESDAPDIDLAAAEATFTGA  
RAREKVEY

15

5 14. DNA sequence according to any of claims 1 to 5 wherein the  
DNA is selected from the group consisting of

(a) the following DNA sequences:

20

10 Seq ID No 119 (>Contig17)

25

TTACGTTACTCATCCTATCTCGGCACCCTGTGTCGGTGATGTCGCTCGCC  
TCGAGCGCGAGCGGGACGACGTCGGCGCCGCGCTCGGTGAGCGCCGCCCG  
GAGGGCGCTCGCGAGATCGCTGGCGACGCCGGCCGGGGCCACGACGAGCC  
ACGTCCCCGCGACGTCGCCGCGTGACGCGGCGCTCACGGGTCTCCATTCTG  
15 ACGCGGTAGCGCCACGCGCCACGGTGCTCTGCTCTCGGCGGCTCCGCCG  
CCACGCCGACAGGGCCGGCATGAGGCTCTCGAGGGCCGAGCGCCGCCCGC  
TGTCGGCGACGTGGAGCGCGTCCGAGAGCGCCGCGACGTCGCCGCGCTCG  
30 ATGGCTCGCCAGAACGCGGTCTCCTCGGCGGACGCTCCCGGCGCCGCGTC  
CTCATCGTCCGACGCGTCCGCTGCGTCGAGCCAGAACCGCTCGCGCTGGA  
20 ACGCGTACGTCGGCAACGTCACGCGGCGCGCCCCGAGCGGAGCGAAGAAC  
35 GCACCCAGTCGATGGCGTGCCCGCGCGCGTGGAGCTCGCCTGCCGAGAG  
GAGGAAGCGCTCGAGGTGCGCTTCGTCGCGGCGGAGCGAGGACACCACGG  
TCGCATCGCCGTCGATCGACGAGAGCGTCTCGTCGAGCGCGACGGTGAGC  
ACGGGGTGAGGGCTGACCTCGACGAAGAAGCGGTGGCCGTCGTCGAGCAG  
40 GGCGCGCGTGCGGTGCTCGAAGCGGACGGTGTGGCGCAGGTTTCGGTACC  
25 AGTGGGCGGCGCCGAGGGCCTCGCCATCAAGCCTCTCGCCCGTCACCGCG  
GAGTAGAGCGGCACGGTCGCCGGGCGCGGCGCGATGCCGTCGAGCGCCTC  
45 CAGCATCGTCCGCTCGATGGCCTCCACGTGGGCGGAGTGGGAGGCGTACT  
CGACGCGGACCTTGCGGGCGAACAGCTGCGCCCCGCTCAGCTCTGCGACG  
30 AGCTCGTCGATAGCGCCGGGGTCTCCGGAGACGAGGGCCGCGTGAGGGCT

50

55

5

230

10

15

20

25

30

35

40

45

50

55

GTTGATCGCCGCTATCGCCAGGCGTTCGCCCAAGGGCGCAAGGCGCGCCT  
CGAGCTCGGCGGTGGTGAGCTCGACGGCGGACATGGCGCCGCGTCCCGCG  
AGCTTCGTAATGGCGCGCGAGCGGAGCGCGACGACCCTGGCGGCGTCTTC  
TAGCGAGAGCGCGCCCGCGACGTACGCGGCCGCGATCTCGCCCTGGCTGT  
GGCCGACGACCGCGTCCGGCGGTGACTCCGGCGGCGCGCCAGGTGGCGGCG  
AGGGCGATCATGACGGCGAACAGCACGGGCTGCACCACGTGACGCGCTC  
GAGCATGGGCGCGGCGTGCGCTTCGTGCGCCGCGAGCACGGCGAGGAGCG  
ACCAGTCGACGTGCGGCGCCAGGGCGCGCTCGCACGCCTCGATCTCGGCC  
CGAAAGGCGGGCGAGGAGGCGAGCAGAGCGCGCGCCATCGATGGCCACTG  
CGAGCCCTGGCCGGGGAAGACGAAGGCGACCTTGCCCGGCGGGAGCGCCT  
CGCCCGCGACCGTTCTGCCCCCGCGCGCCCTCGGGCGAGCGCCGCGAGC  
GCCGAGAGCAGCGCGGCGCGATCGTCTGCCACGACGGCGGCGCGACGCTC  
GAAATGCGACCGCGTGGTCCGAGCGACGCCGCGACGTGACGAGGGCGA  
CGTCCTCGTGCTCGGCGAGGTGCGCGTGAGCTTGCCCGCCTGAGCGCGG  
AGCGCCGCGTCGCTCTTCGCCGAGAGGAGCACCGGCACCGGCGGCGCGAA  
GGGCGCGCGGGCGGGCTCCCCGGCCTGGTCTGCGCCGCGCCGCGCGCG  
GCGCTTCCTCGAGGACCACGTGCGCGTTGGTGCCGAGATCCCGAACGAC  
GACACCGCCGCGCGCCGAGGAGACCCGCCTGGCTTCCACGGTACCTCCTC  
GGTCAAGAGGCGGATCGCGCCGGACGACCAATCGATGTGCTGCGACGGGC  
TCGCGGCGTGAGCGTCTTCGGGAGGACGCCGCTCTGCAGCGCGAGCACC  
ATCTTGATGACGCCCGCGATCCCCGCGGCGGCCTGCGTGTCGCCGAGGTT  
CGACTTTAGGCTCCCGAGCCACAGCGGGCGCTCCTTCGCGTGCGCCGCGC  
CGTACGTGCGGAAGAGCGCGCGCGCCTCGATGGGATCGCCGAGCGTCGTG  
CCGGTTCCGTGCGCCTCGACGGCGTCGACGTCCGCGGGGGCGAGCCCCGC  
GCTCGCGAGCGCGTCCCGATCACGCGCTCTTGCGCGGGGCCGTTTCGGCG  
CCGTGAGCCCTTGCTCTTGCCGTCTTGGTTGACGGCCGATCCGCGCACG  
ATCGCGAGCACGGGGTGCCCGTTCTTCCGGGCGTCCGACAGGCGCTCGAG  
GAGCACTATCCCAGCGCCTTCCGACCAGCCCGCGCCGTTTCGCGTGCGACG  
AGAACGACTTGACCGCCCCGTCCGGCGCGCCCGCGTGCTGCGCGCTGAAC  
TCGCCGAAGATCCCGGGGGTCGCCATCACGGTCACGCCGCGGCGAGCGC

5

231

10

GAGCGAGCACTCGCCTCGACGGATGGCGTGGCAGGCGAGGTGGAGCGCGA  
CGAGCGACGAGCTGCACGCCGTGTCGACGCT

Seq ID No 120 (>Contig18)

15

5 TTTTAGGANCCCCGACGTGCACGATCGGCTCGCCAACCTCGTGGCGCGCC  
GGGACTATTTTACCAGCTCGCGTTGCGCGCCGCGGGGACCTACGTGCGG  
GGCCTCGTCCGCGCCCCGCACGACGGCGCGCGCCCCCGCGTTCGCGCC  
GCGTGGGGCGGCGCTCGTCACGGGCGGGACCGGGGCGCTCGGGGCGCACG  
TTGCCCCGTTGGTTCGCGCGGATCGGCGCCGAGCACATCGTGCTCGCGAGC

20

10 CGCCGCGGAGCCGCGGCCCCCGGCGCGCCGCGCTCGCCGAGGAGCTTTC  
GGTGCTCGGCGCGCGCGTGACGCTGTTGCGTGCGACGTCCCCGATCGTG  
AGGCGGTTCGCGGGGCTCGTGCGCAACGTCAAGGCCGGCGGAGCGACGGTG  
CGCGCCGTGTTCCACGCGGGCGGTGCGATGCACGAGGCGCCGGTCGCCGC  
CATGCGTGTTGAGGAGCTCGCCGACGCGATCGCCGTGAAGGCCCGCGGCG

25

15 CGCAGCACCTCCAAGACGTCTTCGCGCAGCGCCCGCTCAACGCGTTTGTC  
CTCTTCTCGTCAGAAACGGGTGTGTGGGGCGGTGGCCGGCAAGGCGCGTA  
CGCCGCGGCGAACGCGTTCCTCGACGCGCTCGCCGAGGCGCGTCGCGCGG  
ACGGCCTCGCGGCGACCTCGATCGCGTGCGGCGCGTGGGCGGGCGGCGGA  
ATGCTCGCGACCGACGCCGAGCGGCGCTTGAAGCATCGCGGCGTCGCGCC

30

20 GATGGATCCGGAGCTCGCCGTGCGGGCCCTCGCGCACGCGCTCGATCACG  
CCGAGACGTGCCTCGCCGTGCGTGACGTCGACTGGGCGCGCTTCGCCCCG  
TCGTTGCGCTCGGCGCGTCTCGCCCGCTCCTCGACGAGCTCGCGGAGGC  
GCGATCGGCGCTCGACGCGCTGCGCGAGCCACCGGACGACGCGCGCACGG  
CCGCCGGTCCCGAGCCCGCAAGCACGCTGAGGACCACGCTCGCGGCGCTC

35

25 CCGGAGGGCGAGCGCCACCGCCACCTCCTCGCGCTCGTGCGGACGGAGAC  
GGCGGCGGTGCTCGGGCACGCGGACGCGTCGCGCGTTCGAGCCGAACGCG  
GGTTCTTTGACCTCGGGCTCGACTCGCTCATGTCCGTGAGCTCCGCAGG  
CGCGTCCAGCGCGCGACCGGCATCAAGCTCCCGGCGACGCTCGCGTTTGA  
CCACCGACGCCGAGCGCGCTCGCGAGCAAGGTGCTCGCCGCGATCGTCC

40

30 TCCACGACGCGACCCCGCGCGCCTCGCCCGCCGCGGAGCTCGAGCGCCTC

45

50

55



5

10

15

20

25

30

35

40

45

50

55

GAGGGGATGCTCTCGGCGATCTACGCGGACGAAGCGCTCCGCGACGACCT  
CACGGCGCGCCTCCGCGCCTTCCTGGACAAGCGCGCGGTCCGCACCGAAC  
GCCCCGACGACGCCGCGTTTCGCCGAGAAGCTCGGCTCCGCGAGCGCCGAC  
GAACTCATTGCCTGATCGATCAGAAGCTCGGAGATCGCATCGATGTCGA  
5 CCGTTACTAACGACACGCTCACGGAGTACTTGCGGCGCCTCACTCAAGAG  
CTCCACAGGAGCGAGACGCGCCTGCGTGCGACGGAAGAGAGGCGACATGA  
GCCGATCGCCATCGTTCGGCCTCGGGCTCCCCTTCGGGGCGGGATCCACG  
ACCGCGACACGCTCTGGACGTTCTTCGAGGAGGGCCGCGACGCCATCGCG  
CCGATCCTCGCGAGCCGCTGGAACGCGGACGCGACGTACGACCTCGATCC  
10 GGACCGCGTCGGCAAGAGCTACGTGCGCGACGCCGCCATGCTCGATCGCG  
TCGACCTTTTCGACGCCGATTTCTTCGGGATCAGCCCGCGCGAGGCGAAG  
TACGTCGACCCGCGAGCACCGCCTCTTGCTCGAGACGTCTGGCAAGCGCT  
CGAGGACGCGGGGATTGTGCCGGCGTCGCTGCGAGACTCGAAGACCGGCG  
TCTTCGTCGGCACGGGCGCGAGCGACTACGCGTTCTCCAGAGCGATCGC  
15 GACGCCTCGGAGGCGTACGCGTTTCATGGGGATGATCTCGTCGTTTCGCGGC  
GGGCCGCCTCGCGTTCACGCTCGGGCTCCAAGGCCCGCGCTATCGATCG  
ACACGGCGTGCTCTTCGTCGCTCGTCGCGCTCCACCTCGCGTGCCAGTCG  
CTGCGTCAAGGCGAGTGCGACCTCGCGCTCGTCGCGGGTGTGCAGGTCAT  
GTCGTCGCCGAGGTGTTTCGTGCTGCTCTCGCGCACGCGCGCGCTCGCGA  
20 GCGACGGGCGATCGAAGACGTTCTCGGCGAACGCCGACGGCTATGGCCGC  
GGCGAAGGCGTCGTCTGCTGCGCGTTCGAGCGCCTCCGCGACGCGCGCGC  
GAAAGGGCGCCCGATCCTCGCGGTGATCCGCGGCAGCGCGGTGAACCACG  
ACGGCACGTCGAGCGGGATCACGGTCCCGAACGGGCCCGCGCAGCAGAAG  
GTGCTCCGCGCCGCGCTCGACGACGCGCGGCTTGTCGCCCGCGACGTCGA  
25 CGTCGTCGAGTGCCACGGCACGGGGACCTCCATCGGCGATCCCATCGAAG  
TGAACGCGCTCGCCGCCGTCTACGGCGAGGGGCGCCCCAAGGACCGCCCG  
CTGTTCTTGGGCGCGCTGAAGACCAACATCGGGCACCTCGAGTTCGCGTC  
GGGCCTCGCCGGCGTCGCGAAGATGGTCGCCTCCATGCGCCACGCGACCC  
45 TCCCCGCGACGCTGCACACGAGCCCGCTCAACCCGCTCGTCGACTGGGAC  
30 GCGCTCCCCGTGCGCGTCGTGACGCCGCGCGCCCGTGGACGCGCCGCGA

5

10

15

20

25

30

35

40

CGACGGCGCCCCCGGCGCGCCGGCGTCACGGCGATCGTCGAGGAGGCGC  
CCGCCGAGCCCGAGCCACGACGCCGACGCCGCGCCCGCGCTTCCGGCC  
GTGCCCCTTCTCCTCTCGGGCAAGACCGACGAGGCGCTGCGCGCGCAGGC  
AGCGCGCCTCCACGCGCACCTCGCGGGGCGCCCCGACGCGCGGCTCGTCG  
5 ACATCGCCGCGTCGCTCGCGACGACGCGCACGCACTTCGATCGACGCGCG  
GCCGTGTCGCGGCGGATCGCGACGAGCTCCTCGGCGCGCTCGACGCGCT  
CGCGCGCGGCGAGGCAGGCCCGGGGTGCGTCGTCGCGAGCGCGATCCCCG  
CCGGCAGGGTCGTGTTGTTGTTCCCGGCCAAGGCTCGCAGTGGGTGCGG  
ATGGCGCGCGCGCTCCTCGCGTCGTCGGTGGTCTTCCGCGACGAGATCGC  
10 GGCTGCGAGCGCGCGCTCGCGCCGACGTCGCTGGTCGCTCGGCGCCG  
TTCTCCGGGGCGACGGCGACGAGGCGACGCTCCTCGGCCGCGTCGACGTC  
GTGCAGCCGGTCCTCTTCGCCGTCATGGTCGCCCTCGCCGCGCTCTGGCG  
CTCGATCGGCGTCACGCCGACGCCGTCGTCGGGCACAGCCAAGGCGAGA  
TCGCCGCCGCCTACGTCGCCGGCGCCCTCTCGCTCGAAGACGCCGCCAAG  
15 GTCGTCGCGCTGCGCGCACGAGCGCTCACGAAGATCGCGGGGCGCGGGGC  
GATGGCCGCCGTCGAGCTCGGCGCACGCGACACCGAGGCGCGCCTCGCGC  
CGTTCGGCGACGCCATCGCGATCGCGGCGATCAACAGCCCGCGCGCCACG  
CTCGTCGCGGGCGACACGGACGCGATCGACGCGCTCGTCCGCGACCTCGA  
GGCCGCGCAGATCTTCGCGCGGAAGGTGCGTGTCGACTACGCGTCGCACT  
20 CGGCGCACGTCGAGGCGATCGAGCGCGAGCTCCTCGCGGATCTCGCGGGG  
ATCGAACC GCGCGCGGGGCGCTGTGCCGCTTTACTCCGCGGTGACGGGCGC  
GAAGCTCGACGGGAACCGCCTCGACCCCGCGCATTGGTTCCGGAACCTGC  
GCTCGACAAAAA ACTTTGAGGACGCCACGCGCGCGCTCCACGACGACGGC  
CGCCGGGTATCCTCATNATCNNGGGCGTNCAGAGGAGTCGGTATTNCCCC  
25 CCCCCGCCTTNCCCG,

or their complementary strands,

45

50

55

5

10

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

15

5 (c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

20

10 (d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

25

15 15. Peptide encoded by a DNA sequence according to claim 14 selected from the group consisting of

30

Seq ID No 121

35

>Contig17\_001 828 amino acids MW=86259 D pI=5.60 numambig=1  
MTVMATPGIFGEFSAQHAGAPDGRCKSFSSHANGAGWSEGAGIVLLERLSDARKNGHPVL  
20 AIVRGSAVNQDGKSQGLTAPNGPAQERVIRDALASAGLAPADVDAVEAHGTGTTLGDP  
ARALFATYGAAHAKERPLWLGSLKSNLGHGTQAAAGIGGVKIMVLALQSGVLPRTLHAASP  
SQHIDWSSGAIRLLTEVPWKPGGSPRAAVSSFGISGTNAHVVLEEAPRAAAGDDQAGE  
PARAPFAPPVPVLLSAKSDAALRAQAGKLHAHLAEHEDVALVDVAASLATRSHFERRAA  
VVADDRAALLSALAALAEGRAGAGTVAGEALPPGKVAFFVPGQGSQWPSMARALLASSPA  
40 FRAEIEACERALAPHVDWSSLAVLGGDEAHAAPMLERVDVVPVLFVMIALAATWRAAG  
25 VTPDAVVGHSSQGEIAAAYVAGALSLEDAARVVALRSRAITKLAGRGAMSAVELTTAELEA  
RLAPLGERLAIAAINSPHAALVSGDPGAIDELVAELSGAQLFARKVRVEYASHSAHVEAI  
45 ERTMLEALDGIAPRPATVPLYSAVTGERLDGEALGAHWYRNLRHTVRFEHATRALLDDG  
HRFFVEVSPHPVLTVALDETLSSIDGDTVSSLRDEGDLERFLLSAGELHARGHAIDW  
30 GAFFAPLGARRVTLPTYAFQRRERFWLDAGDASDDEDAAPGASAEETAFWRAIERGDVAAL

50

55

235

5

SDALHVADSGRRSALESMPALSAWRRSRREQSTVGAWRYRVEWRPVSAASRGDVAGTWL  
VVAPAGVASDLASALAAALTERGADVPLALEASDITDTGCRDRMSNVX

10

Seq ID No 122

5 >Contig18\_002 502 amino acids MW=53019 D pI=6.83 numambig=1  
15 FRXPDVHDLRLANLVARRDYFYQLALRAAGTYVRGLVRAPHDGARPPAFAPRGAALVTGGT  
GALGAHVARWFARIGAHEHIVLASRRGAAAPGAAALAEELSVLGARVTLVACDVPDREAVA  
GLVRNVKAGGATVRAVFHAGGAMHEAPVAAMRVEELADAIKARGAQHLQDVFAQRPLN  
20 AFVLFSSSETGVWGGGRQGAYAAANAFDALAEARRADGLAATSIANGAWAGGGMLATDAE  
10 RRLKHRGVAPMDPELAVAALAHALDHAETCLAVADVWARFAPSFASARPRPLLDELAEA  
RSALDALREPPDDARTAAGPEPASTLRITTLAALPEGERHRHLLALVRTETA AVLGHADAS  
RVEPNRGFFDLGLDSLMSVELRRRVQRATGIKLPATLAFDHPTPSALASKVLAAIVLHDA  
25 TPRASPAAELERLEGMLSAIYADEALRDDLTARLRAFLDKRAVRTERPDAAFAEKLGS  
SADELIRLIDQKLGDRIDVDY\*

15

Seq ID No 123

30 >Contig18\_010 840 amino acids MW=88062 D pI=5.74 numambig=6  
MSTVTNDTLTEYLRRLTQELHRSETRLRATEERRHEPIAIVGLGLPFRGGIHDRDTLWTF  
LEEGRDAIAPILASRWADATYDLDFDAVGKSYVRDAAMLDRVDLFDADFFGISPREAKY  
20 VDPQHRLLETSWQALEAGIVPASLRDSKTGVFVGTGASDYAFLOSDRDASEAYAFMGM  
35 ISSFAAGRLAFTLGLQGPALSIDTACSSSLVALHLACQSLRQGECDLALVAGVQVMSSPE  
VFVLLSRTRALASDGRSKTFSANADGYGRGEGVVLAVERLRDARAKGRPILAVIRGSAV  
NHDGTSSGITVPNGPAQQKVLRAALDDARLVPADVVECHGTGTSIGDPIEVNALAAVY  
40 GEGRPKDRPLFLGALKTNIGHLEFASGLAGVAKMVASMRHATLPATLHTSPLNPLVDWDA  
25 LPVRVVDAAARPWTRRDDGAPRRAGVTAIVEEAPAEPEPTTPDAAPALPAVPVLLSGKTDE  
ALRAQAARLHAHLAGRPDARLVDIAASLATRTHFDRRAAVVAADRDELLGALDALARGE  
AGPGSVVASAIPAGR VVFVPGQGSQWVG MARALLASSVFRDEIAACERALAPHVAWSL  
45 GAVLRGDGDEATLLGRVDVVPVLFVAVMVALAALWRSIGVTPDAVVGHSSQGEIAAAAYVAG  
ALSLEDAAKVVALRARALTKIAGRGMAA VELGARDTEARLAPFGDAIAIAAINS PRATL  
30 VAGDTDAIDALVRDLEAAQIFARKVRVDYASHSAHVEAIERELLADLAGIEPRAGAVPLY

50

55

5

236

SAVTGAKLDGNRLDPAHWFRNLRSTKNFEDATRALHDDGRRVSSXSXAXRGVGIXPPRLX

10

X

16. Recombinant expression vector which comprises a DNA-  
5 sequence according to any of claims 1 to 10, 12 and 14.

15

17. Procaryotic or eucaryotic cell which has been transfected  
or transformed with a DNA-sequence according to any of claims 1  
to 10, 12 and 14 or with a recombinant expression vector ac-  
20 cording to claim 16.

20

18. Cell according to claim 17, wherein the cell is derived  
25 from myxobacteria.

25

19. Cell according to claim 17, wherein the cell is derived  
15 from a Sorangium strain.

30

20. Cell according to claim 17, wherein the cell is derived  
from Sorangium cellulosum.

35

21. Cell according to claim 17, wherein the cell is derived  
20 from a Streptomyces strain.

40

22. Cell according to claim 17, wherein the cell is derived  
25 from Escherichia coli.

45

23. Process for an enzymatic biosynthesis, mutasynthesis or  
partial synthesis of polyketide or heteropolyketide compounds,  
wherein a cell according to any of claims 17 to 22 is culti-

50

55

5

vated in a suitable culture medium and the polyketide or heteropolyketide compound is isolated from the medium.

10

24. Process according to claim 23, wherein the polyketide or heteropolyketide compound is an epothilone.

15

20

25

30

35

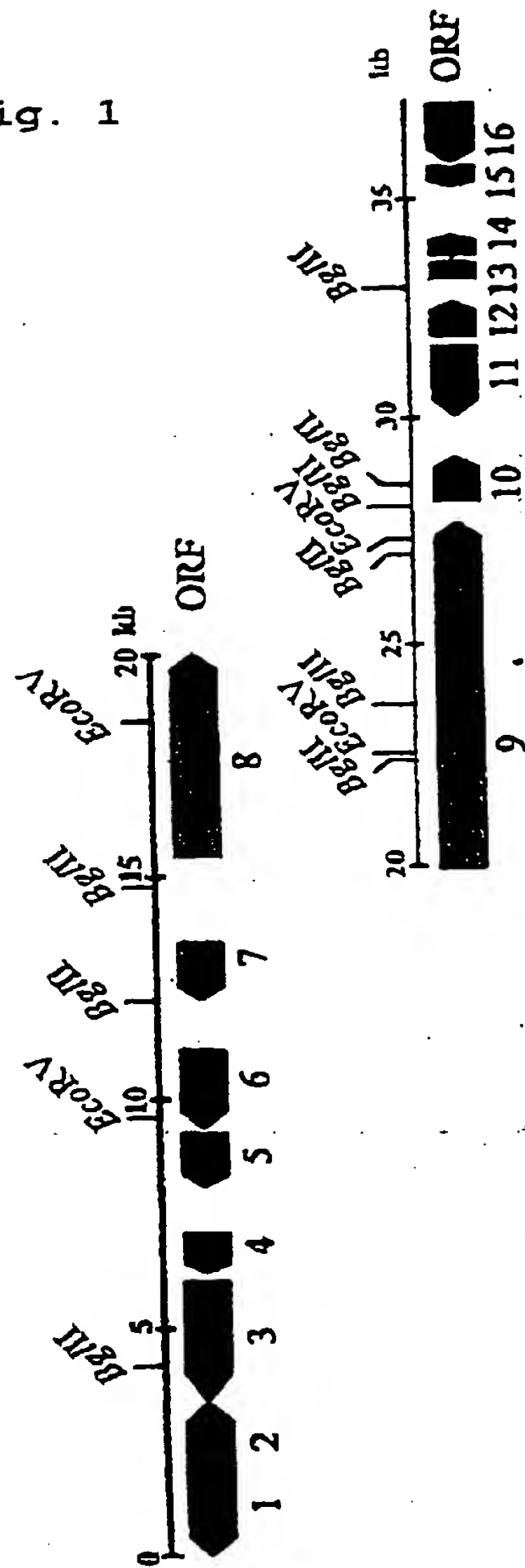
40

45

50

55

Fig. 1



Open reading frames found on pEPOcos6 region.  
pEPOcos6\_annotated  
47713 bp

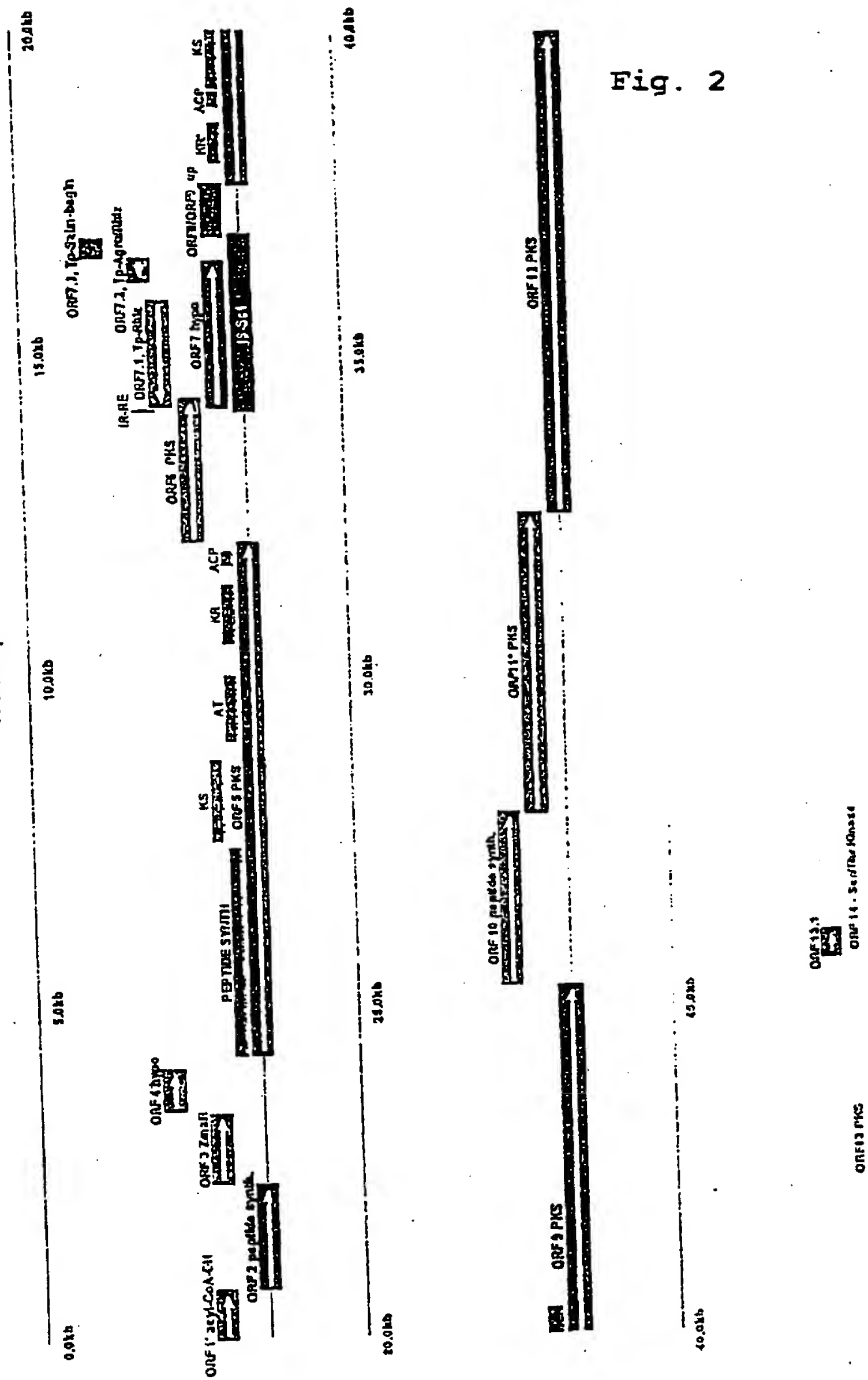


Fig. 2